Supporting Information

Structure-aware fragment assignment for interpreting tandem mass spectrometry of modified and cyclic peptides

Erckes, V.; Misiek, A; Streuli, A.; Steuer, C.*

ETH Zurich, Institute of Pharmaceutical Sciences, Laboratory of Pharmaceutical Analytics, Zurich, Switzerland

* Corresponding author: christian.steuer@pharma.ethz.ch

ORCID

VE: https://orcid.org/0000-0002-9650-4160

AS: https://orcid.org/0000-0002-0025-8023

CS: https://orcid.org/0000-0002-6102-3367

Table of Contents

S1 Peptide synthesis

Chemicals

All Fmoc and side chain protected amino acid building blocks, preloaded Wang resin (0.57 mmol/g) and 2-(6-Chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HCTU) were purchased from Gyros Protein Technologies (Uppsala, Sweden). Dimethylformamide (DMF; EMPLURA Supelco) and diethyl ether (Emsure Supelco, Ph. Eur.) were purchased from Merck (Darmstadt, Germany). Pyrrolidine (99+%) and N-methylmorpholine (NMM; 99+%) were purchased from Thermo Scientific (MA, USA). Trifluoroacetic acid (TFA; 99.5%) was purchased from Apollo Scientific (Stockport, UK). Triisopropylsilane (TIS) was purchased from Tokyo Chemical Industry (Tokyo, Japan). N,N-diisopropylethylamine (DIPEA) was purchased from Carl Roth (Karlsruhe, Germany). Dichloromethane (DCM; >99.8%) and formic acid (FA; 98.0-100%) were purchased from Sigma Aldrich (Buchs, Switzerland). Nanopure water was used from an in-house ELGA Purelab purification system (VWS, Villmergen, Switzerland).

Peptide synthesis

Peptide synthesis of linear peptides was conducted on a PurePep Chorus peptide synthesizer (Gyros Protein Technologies, Uppsala, Sweden). Wang resin with a loading equivalent to 0.1-0.2 mmol (1 eq.) was swollen in 3 mL DMF at room temperature for 15 min. Fmoc deprotection was conducted twice with 2.5 mL 20% (v/v) pyrrolidine in DMF for 5 min.[1] Protected amino acids (5 eq.) were activated with NMM (10 eq.) and HCTU (5 eq.) in 7 mL DMF.[2] The coupling reaction was performed for 10 min at 60 °C (except His 50 °C).[3] All steps were performed while shaking. After each deprotection and coupling step, the resin was rinsed 3 times with DMF. Deprotection and coupling cycles were repeated until the desired sequence was synthesized. The subsequent steps were conducted manually. Before cleavage, the resin was washed three times with DCM. Peptide cleavage from the resin and simultaneous removal of side chain protecting groups was performed in 3 mL of a mixture of TFA, TIS and H_2O with a ratio of 95/2.5/2.5 (v/v) for 2 h at room temperature. The peptide was precipitated from the cleavage mixture by addition of ice-cold diethyl ether. The precipitate was washed and collected by centrifugation (10 min, 10’000 rpm, -10 °C).

S2 RPLC-MS/MS data

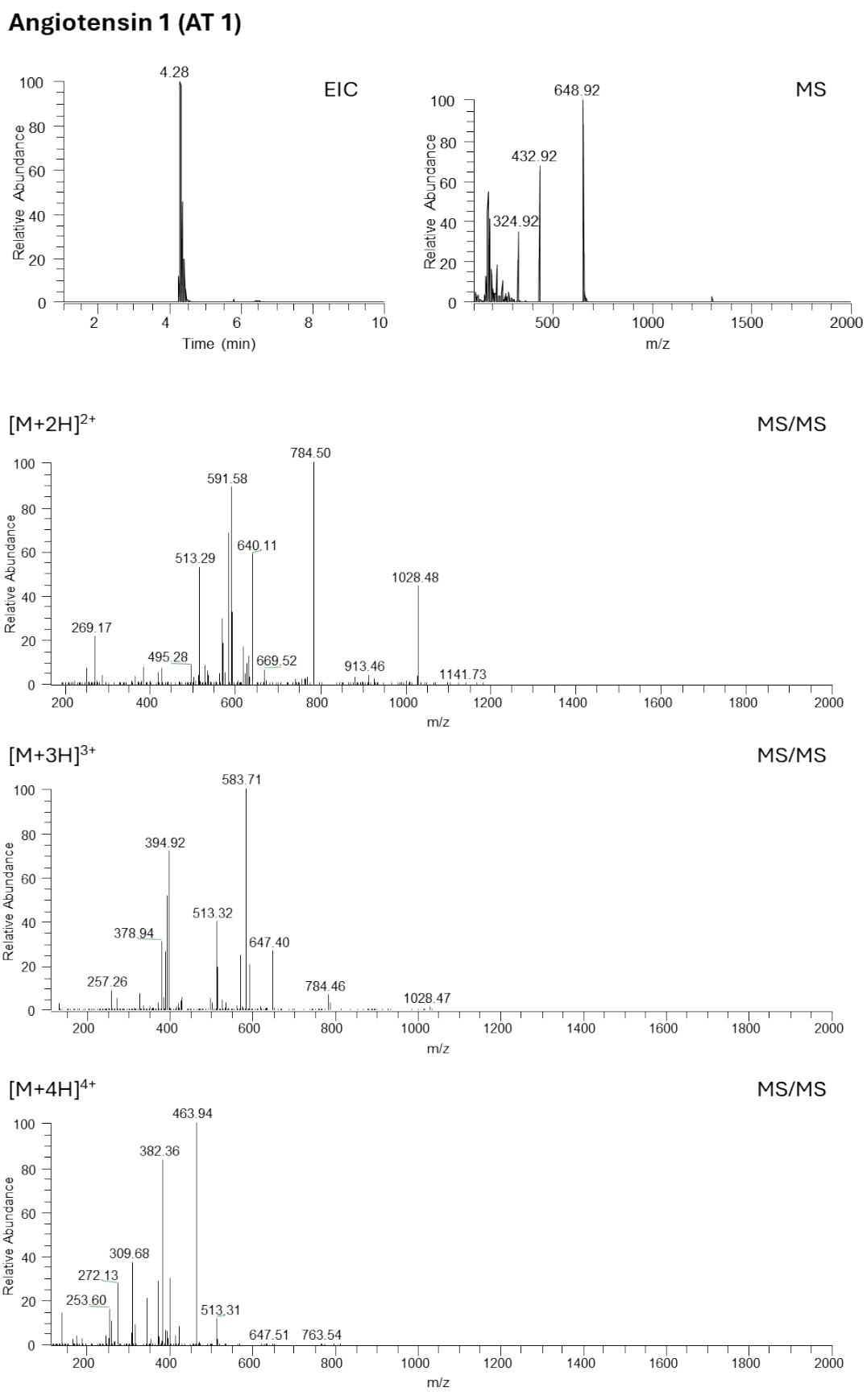


Figure S2-1: EIC, MS and MS/MS data AT1

Angiotensin 2 (AT 2)

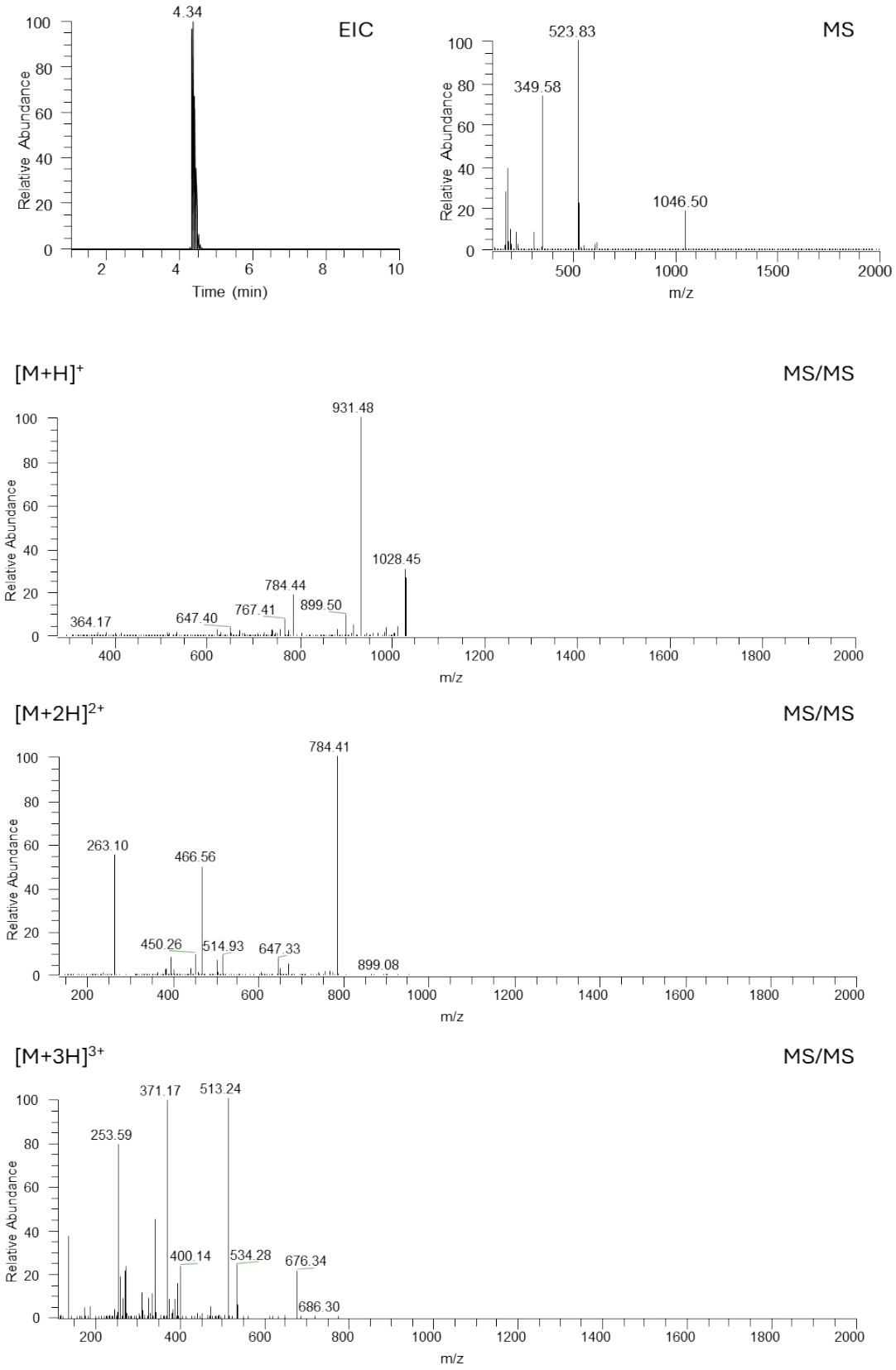


Figure S2-2: EIC, MS and MS/MS data AT2

Angiotensin 3 (AT 3)

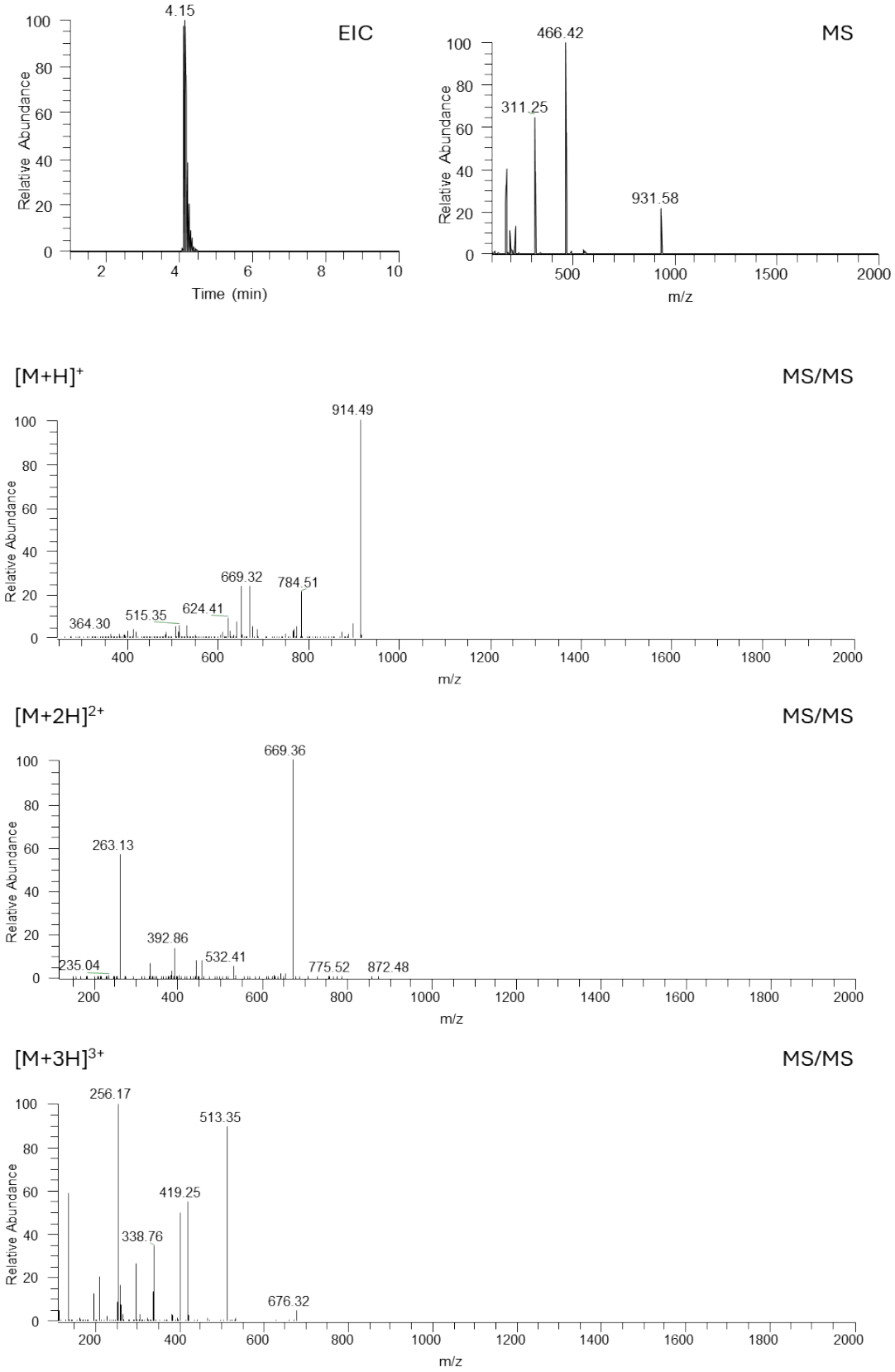


Figure S2-3: EIC, MS and MS/MS data AT3

Angiotensin 4 (AT 4)

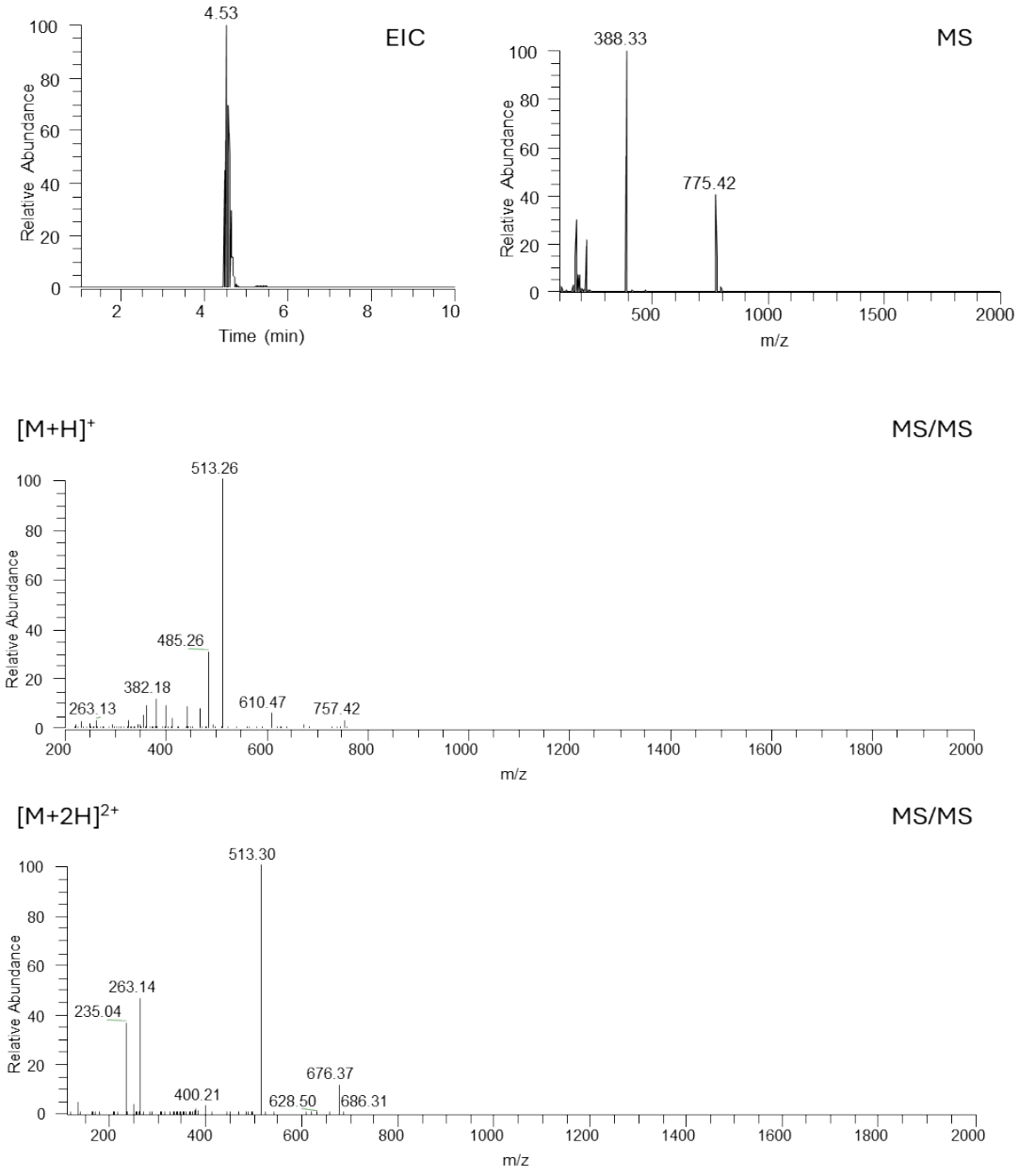


Figure S2-4: EIC, MS and MS/MS data AT4

Liraglutide (LGL)

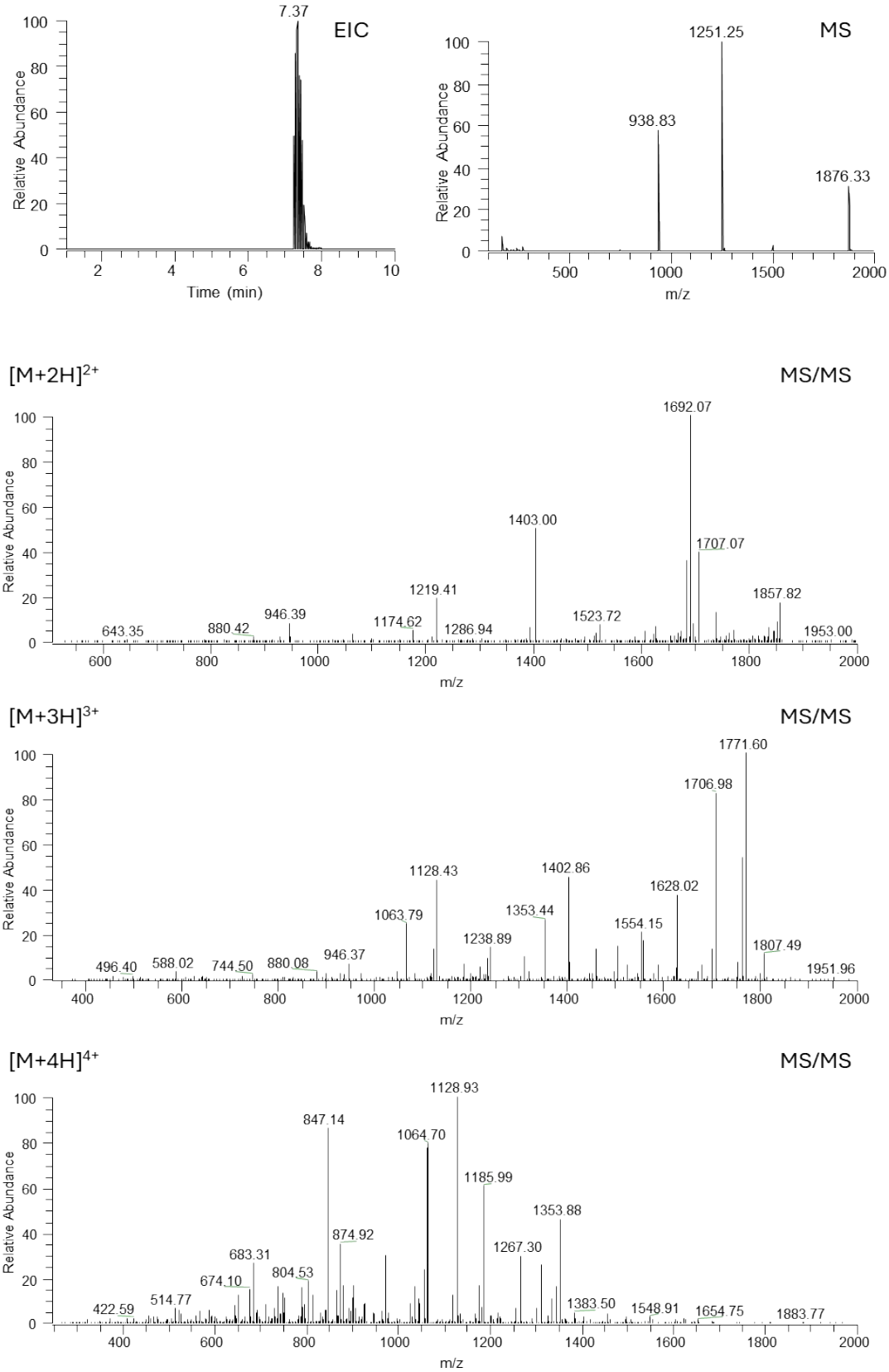


Figure S2-5: EIC, MS and MS/MS data LGL

Semaglutide (SGL)

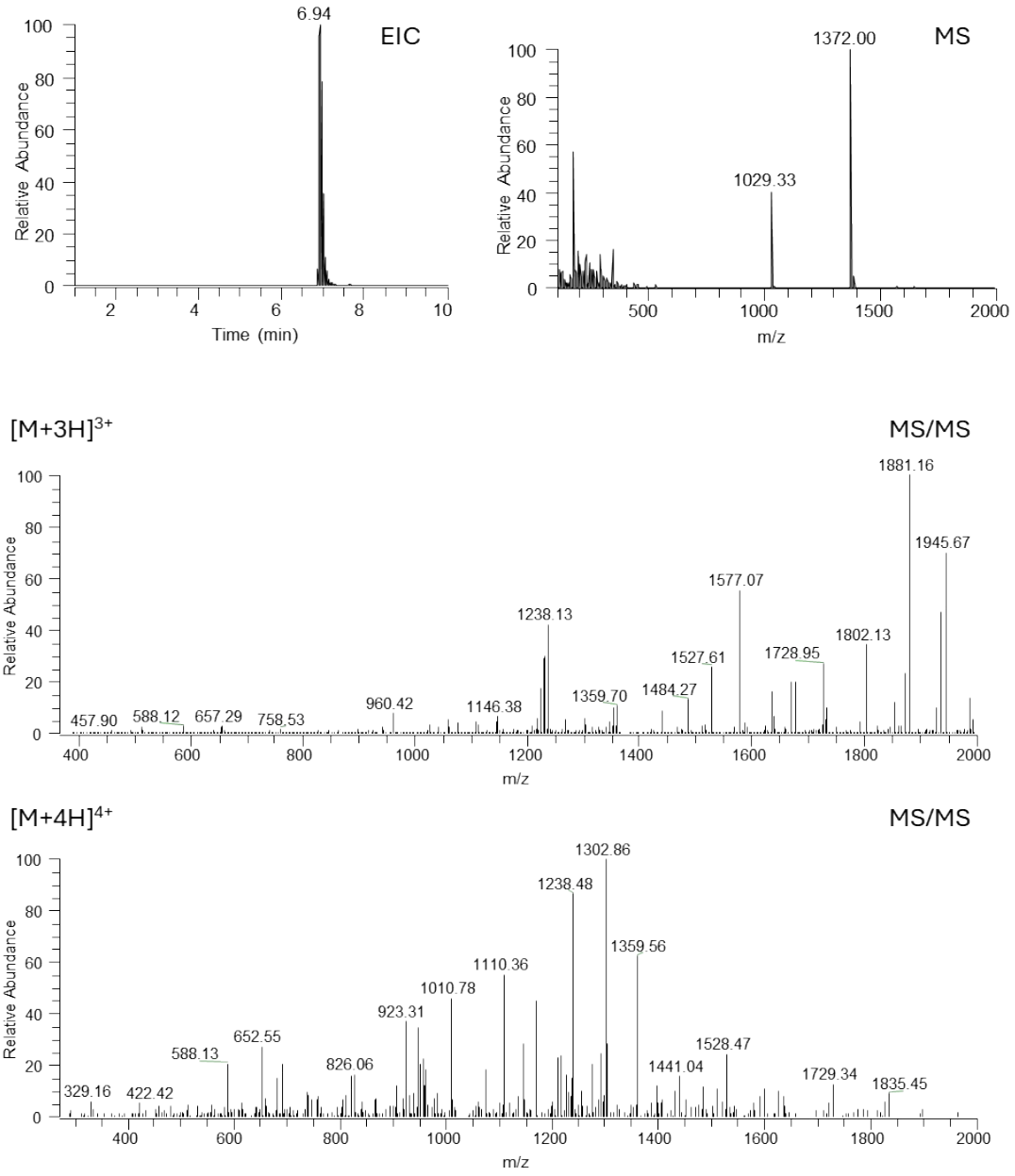


Figure S2-6: EIC, MS and MS/MS data SGL

Cyclosporin (CSA)

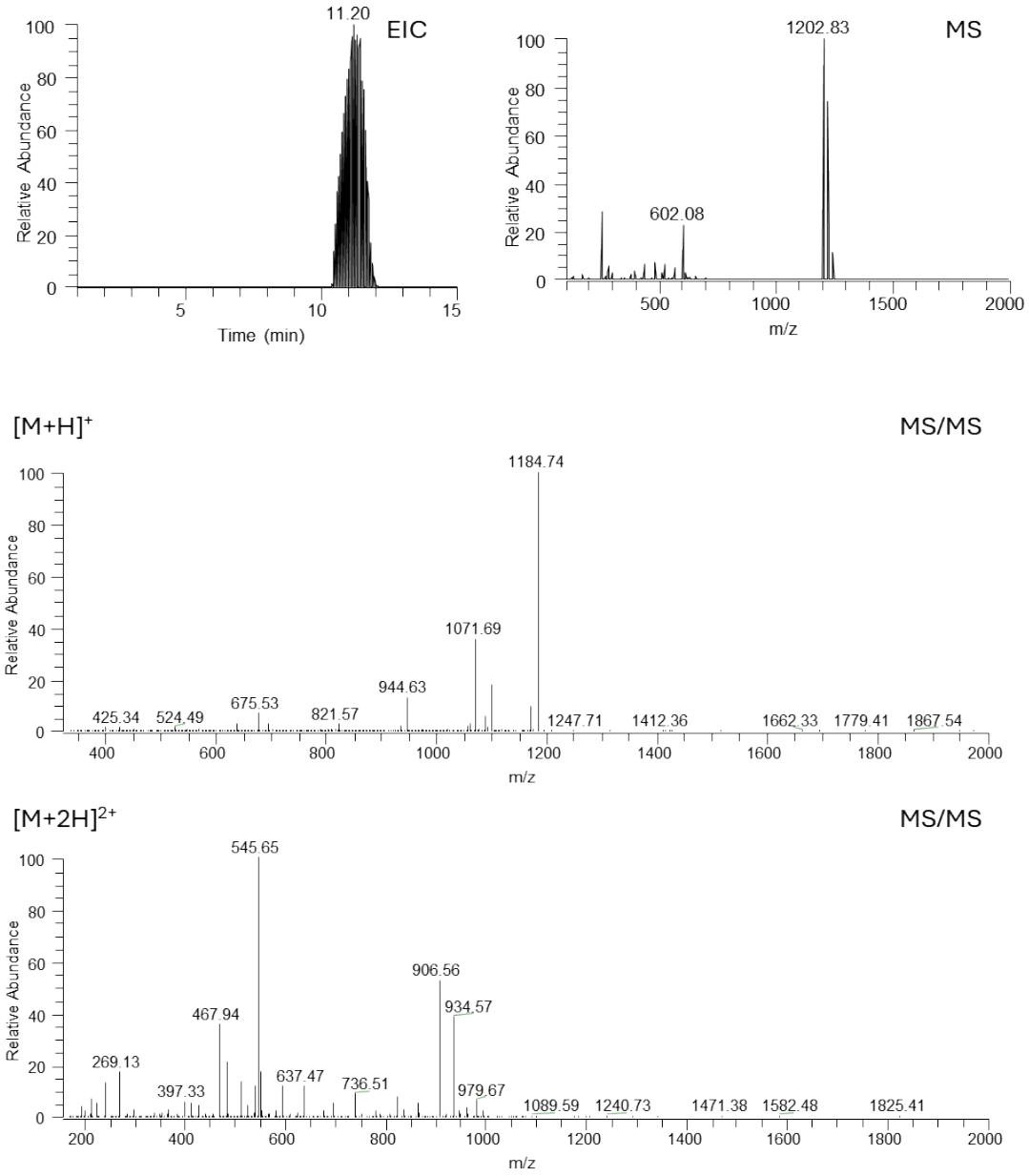
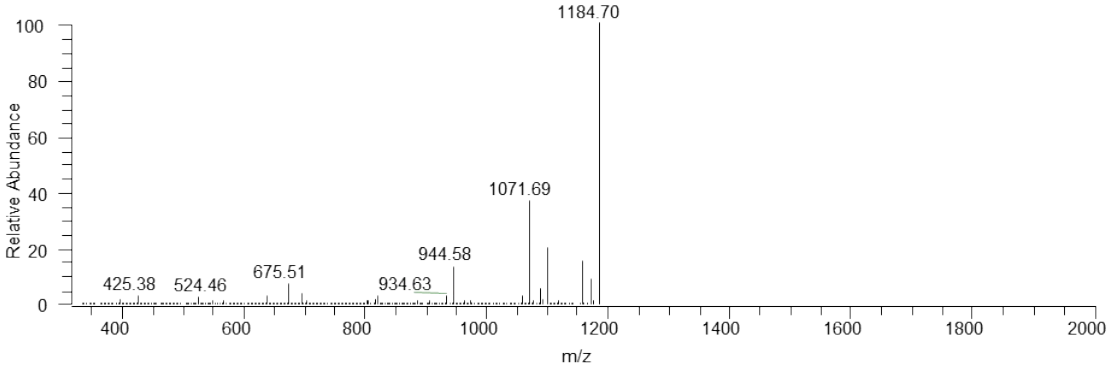


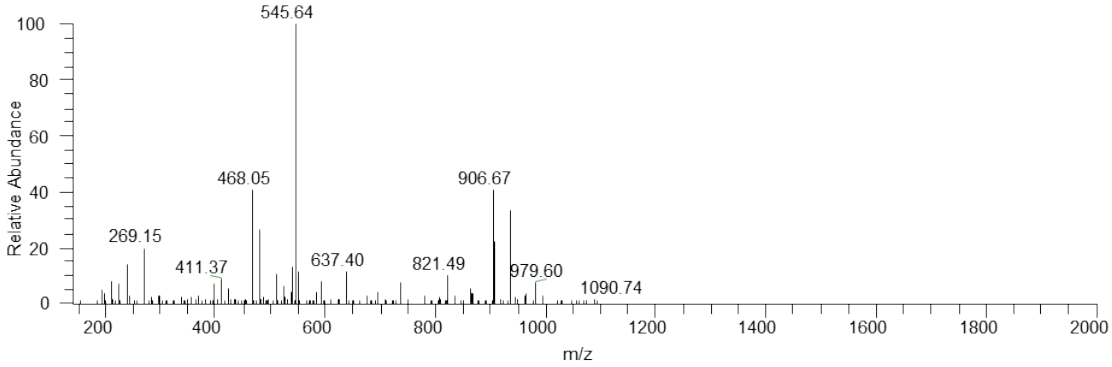
Figure S2-7: EIC, MS and MS/MS data CSA

Cyclosporin (CSA) MS^n

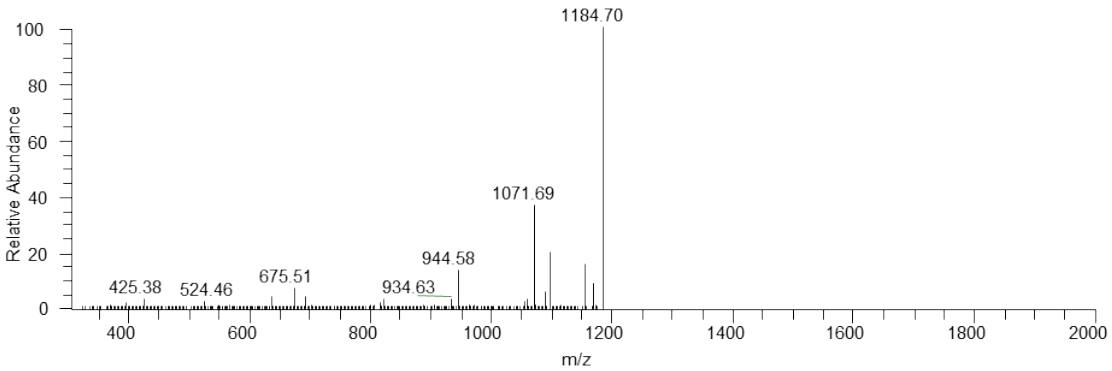
$[M+H]^+$ MS^3



$[M+2H]^{2+}$ MS^3



$[M+H]^+$ MS^4



$[M+2H]^{2+}$ MS^4

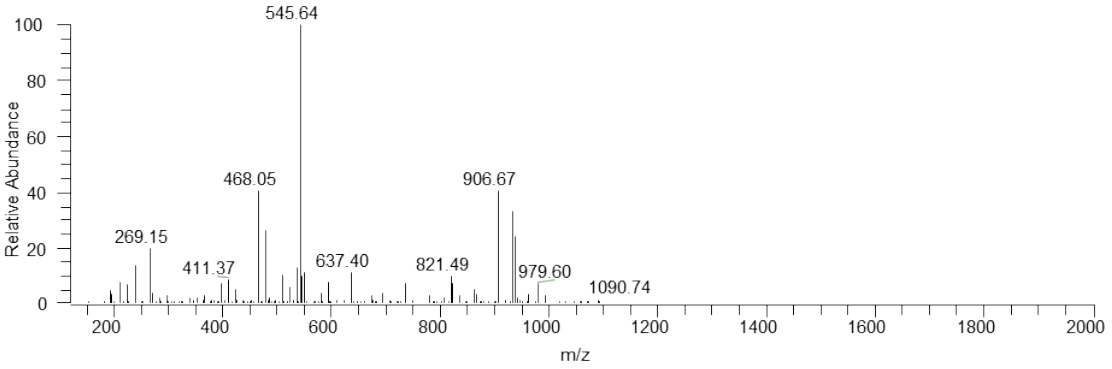


Figure S2-8: MS^n data CSA

Oxytocin (OXY)

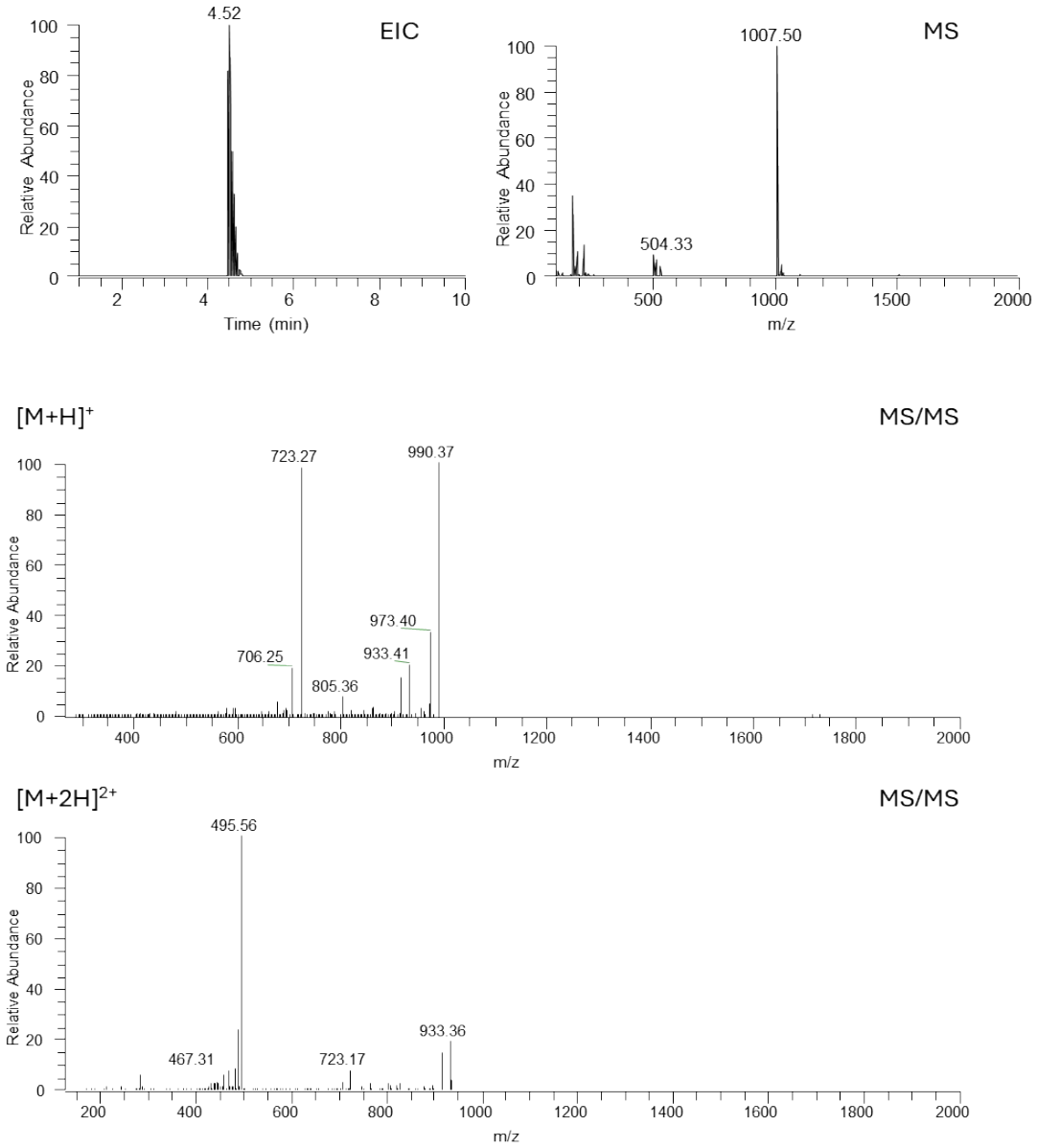


Figure S2-9: EIC, MS and MS/MS data OXY

Somatostatin (SMT)

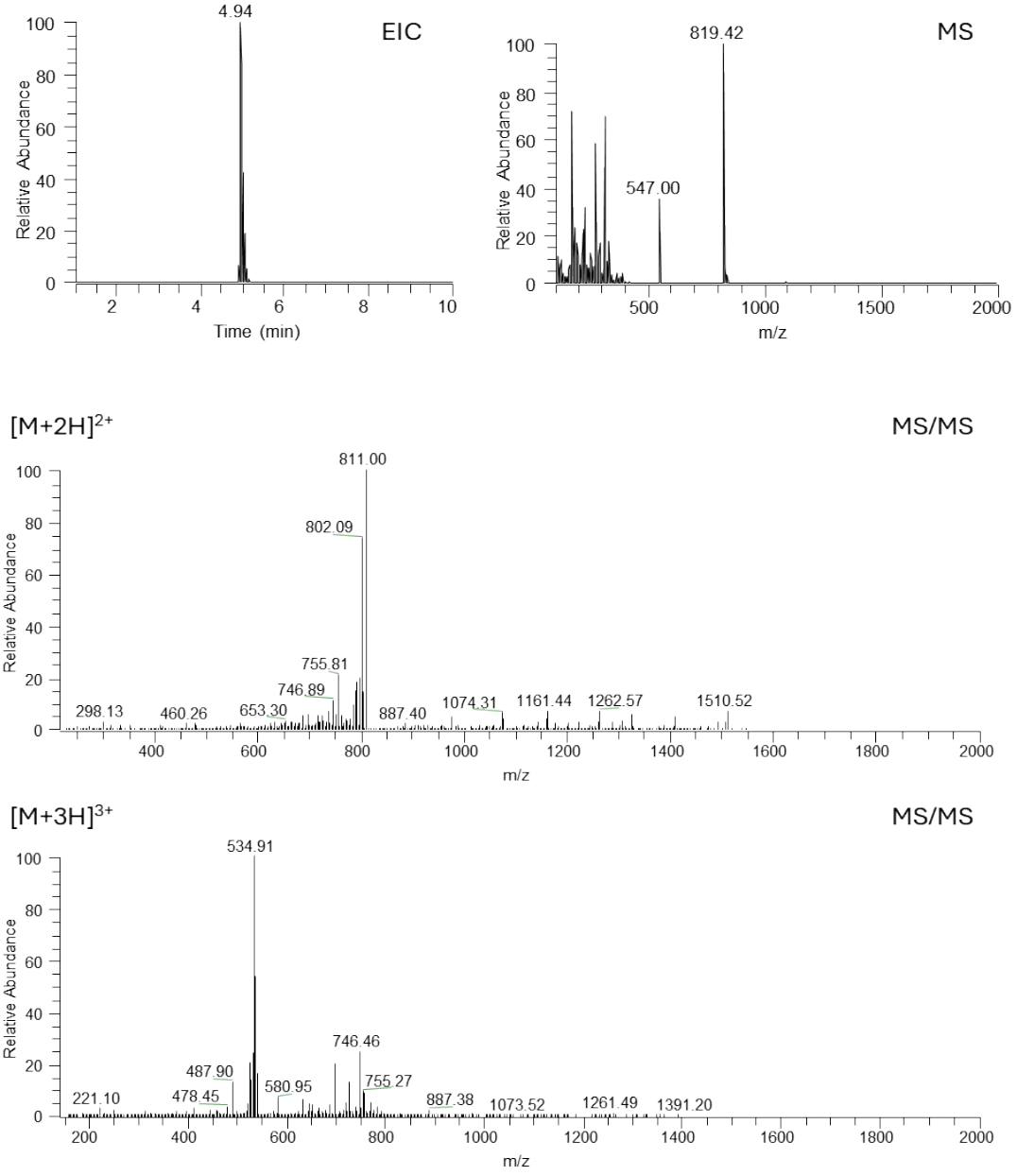
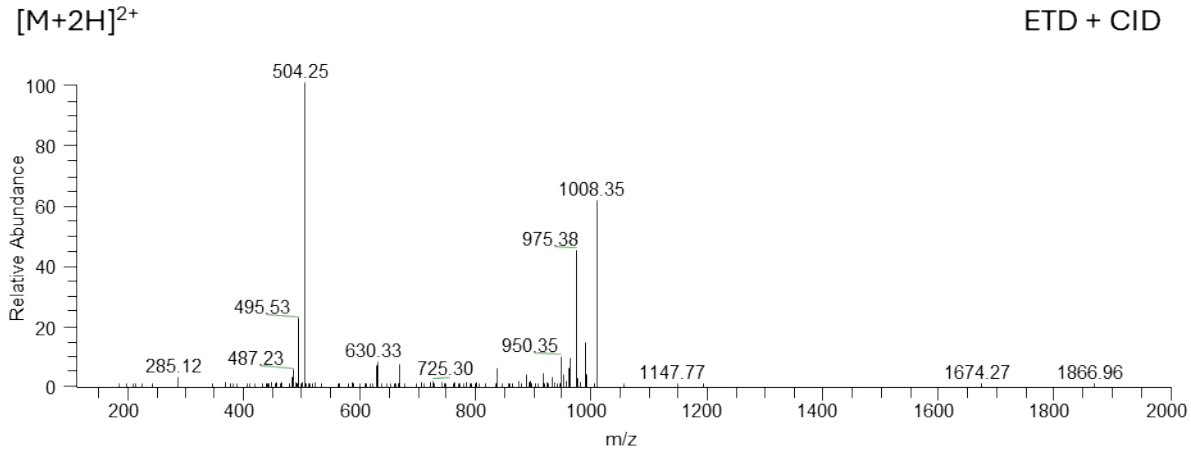


Figure S2-10: EIC, MS and MS/MS data SMT

Oxytocin (OXY) ETD + CID



Somatostatin (SMT) ETD + CID

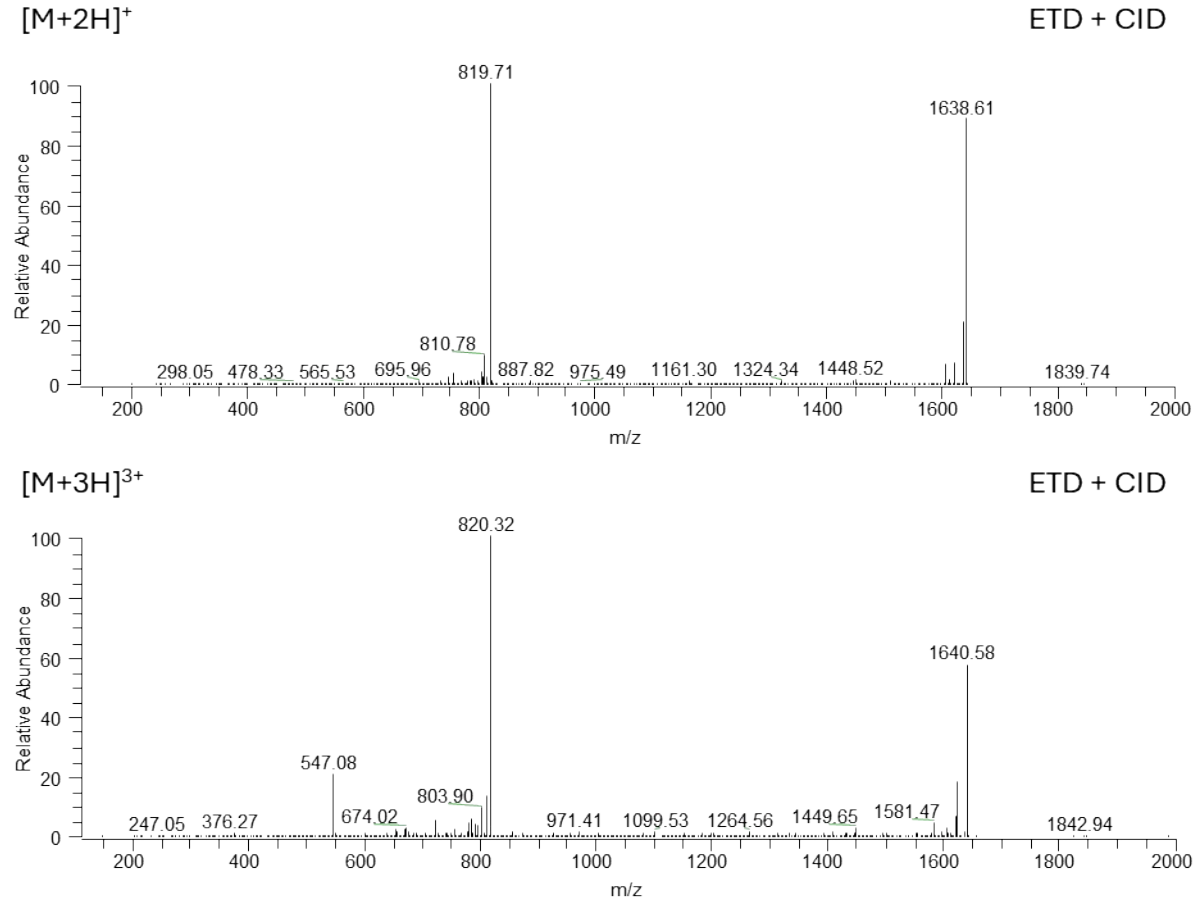


Figure S2-11: MS/MS ETD + CID of OXY and SMT

S3 Assigned MS/MS data

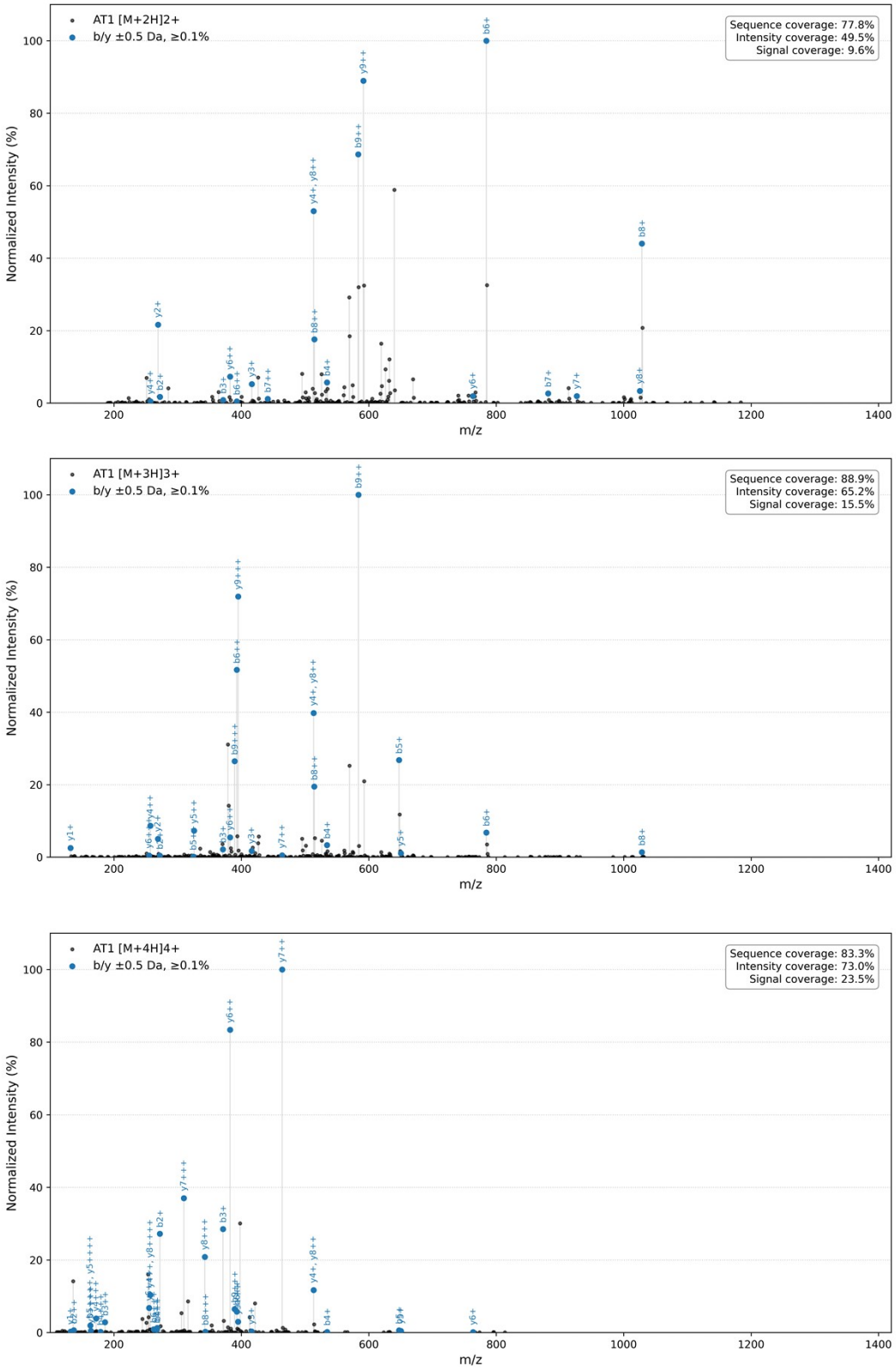


Figure S3-1: MS/MS data assigned AT1

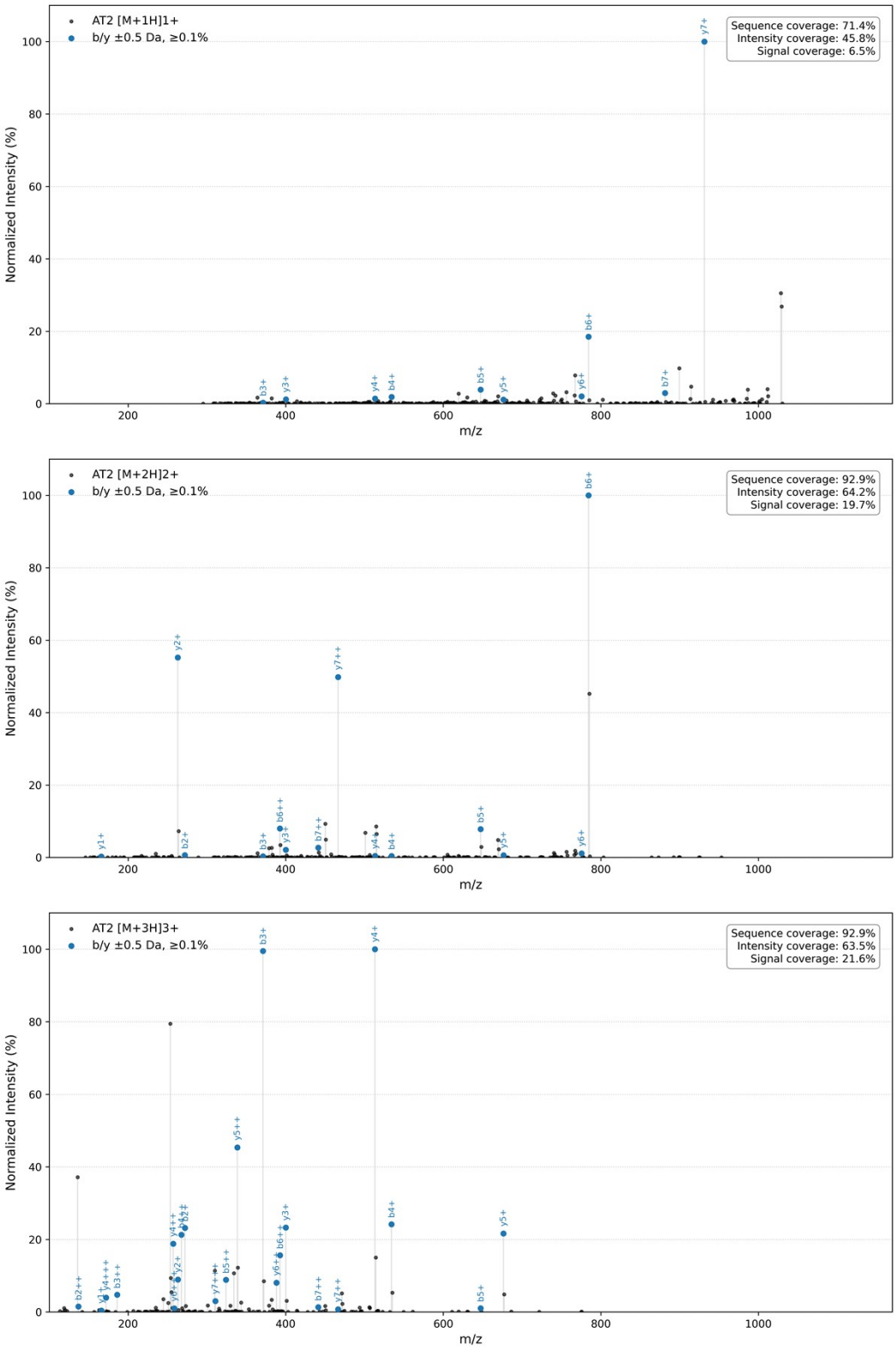


Figure S3-2: MS/MS data assigned AT2

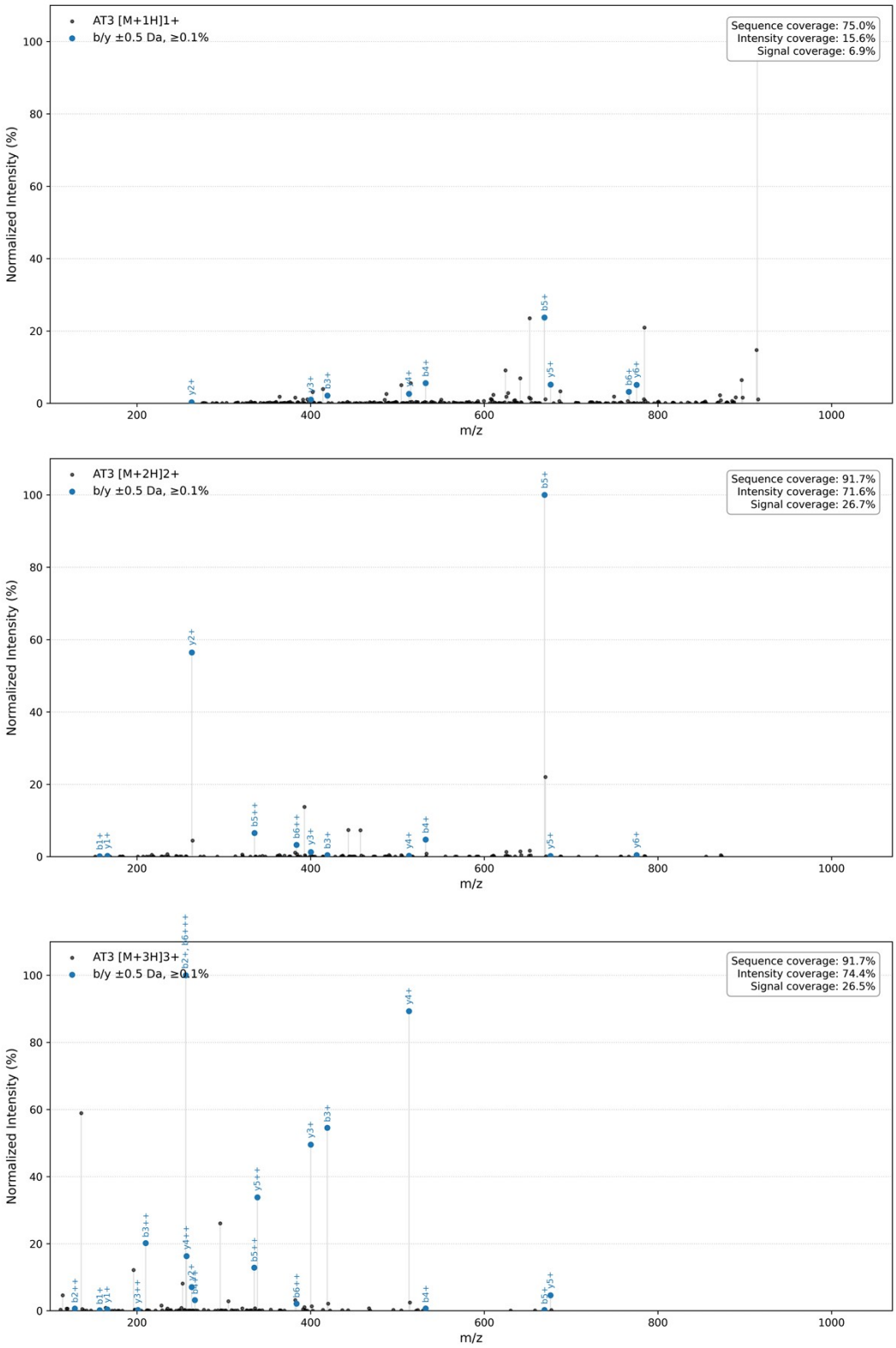


Figure S3-3: MS/MS data assigned AT3

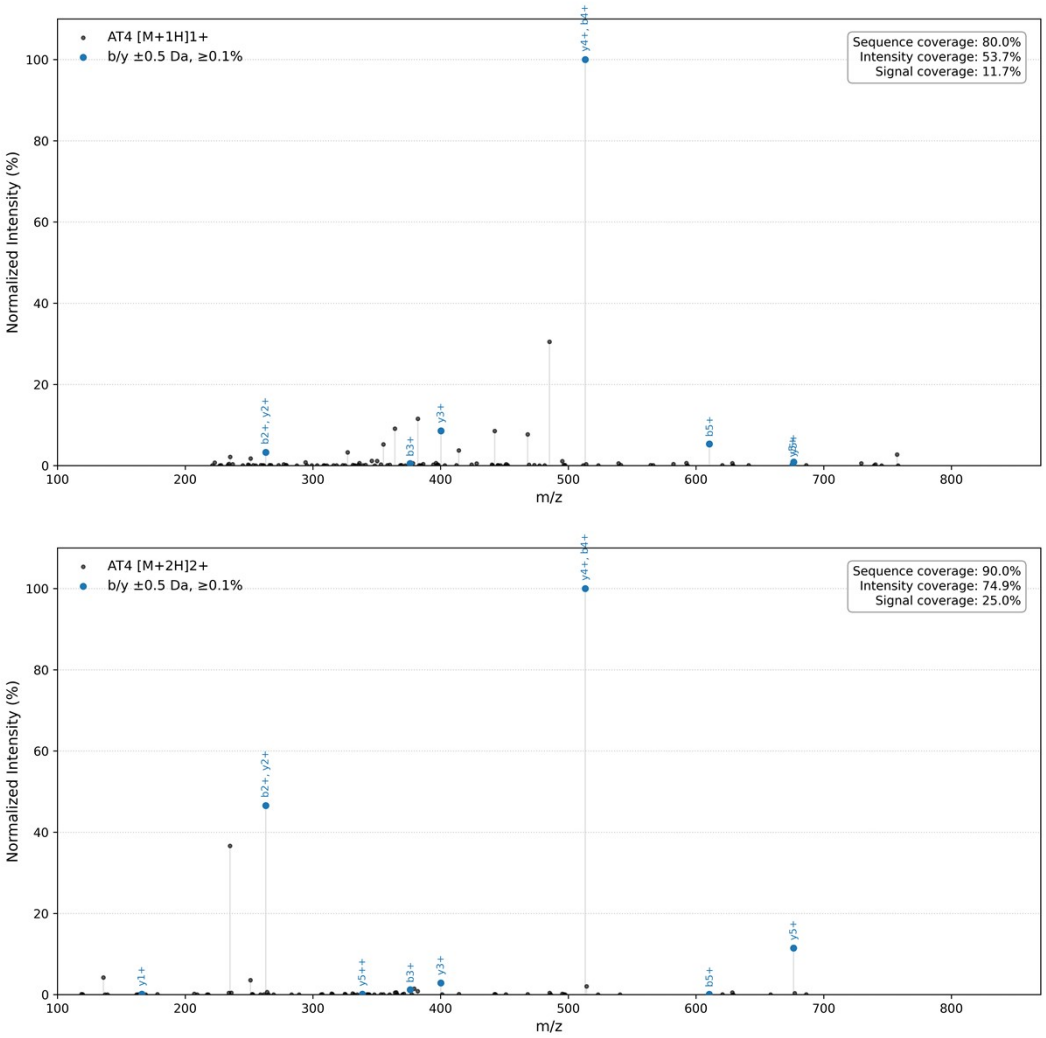


Figure S3-4: MS/MS data assigned AT4

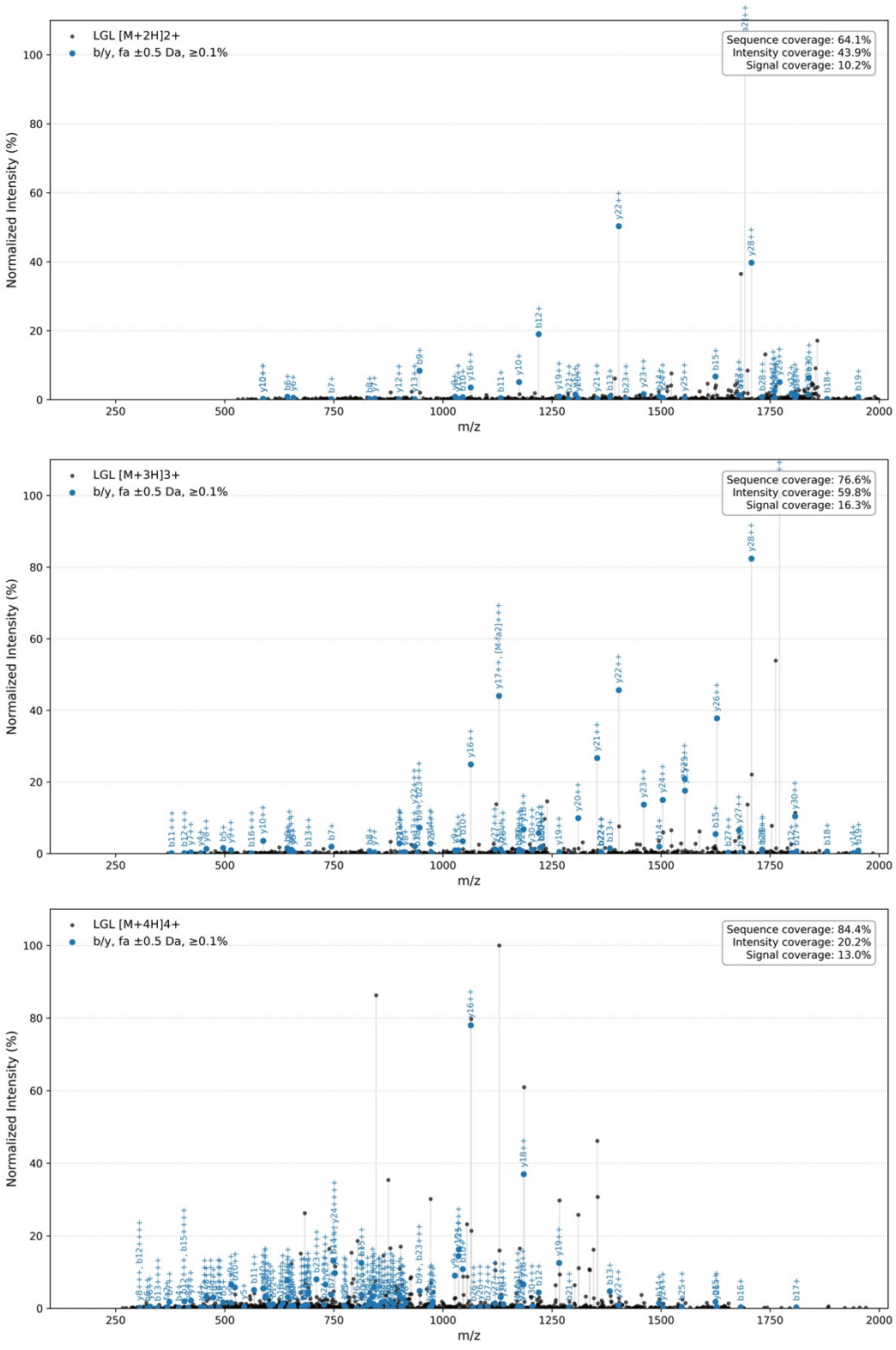


Figure S3-5: MS/MS data assigned LGL

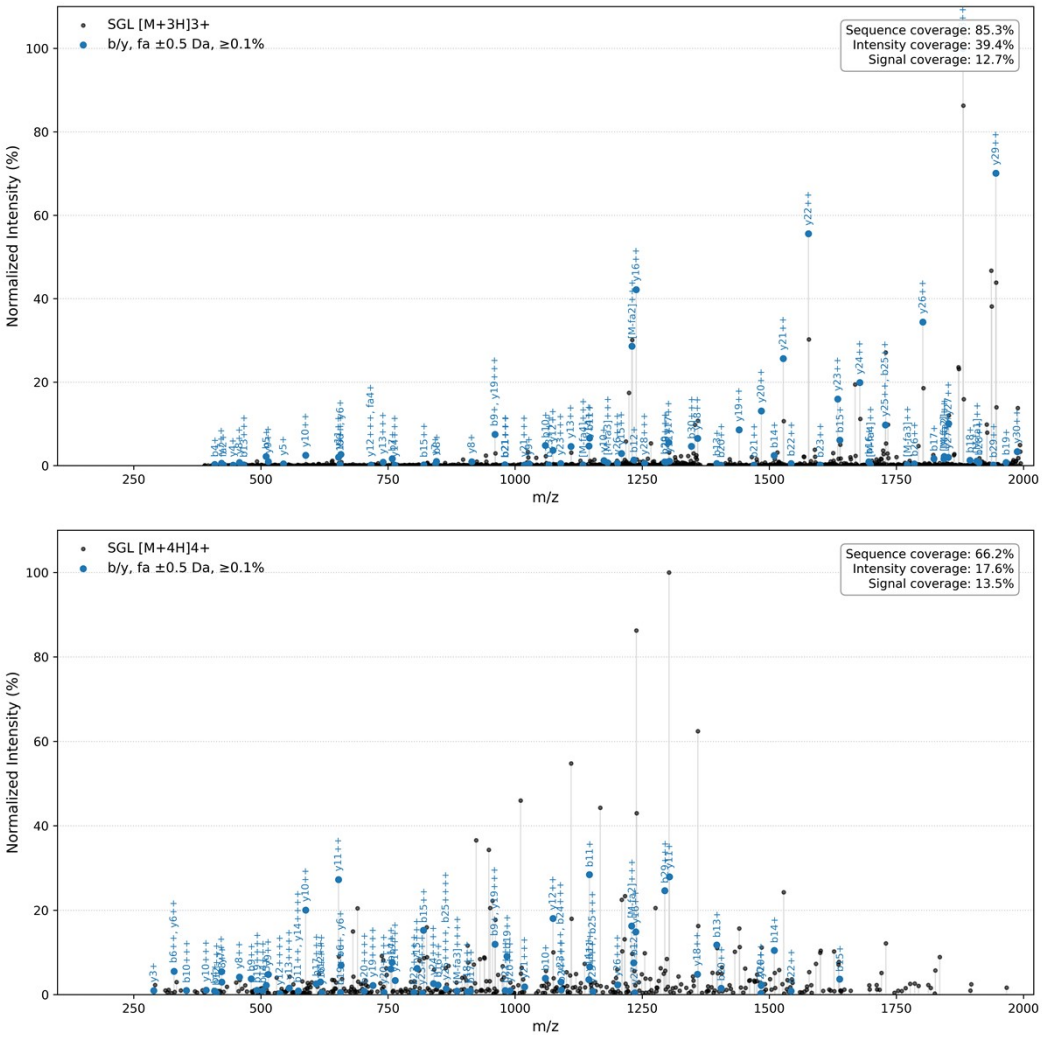


Figure S3-6: MS/MS data assigned SGL

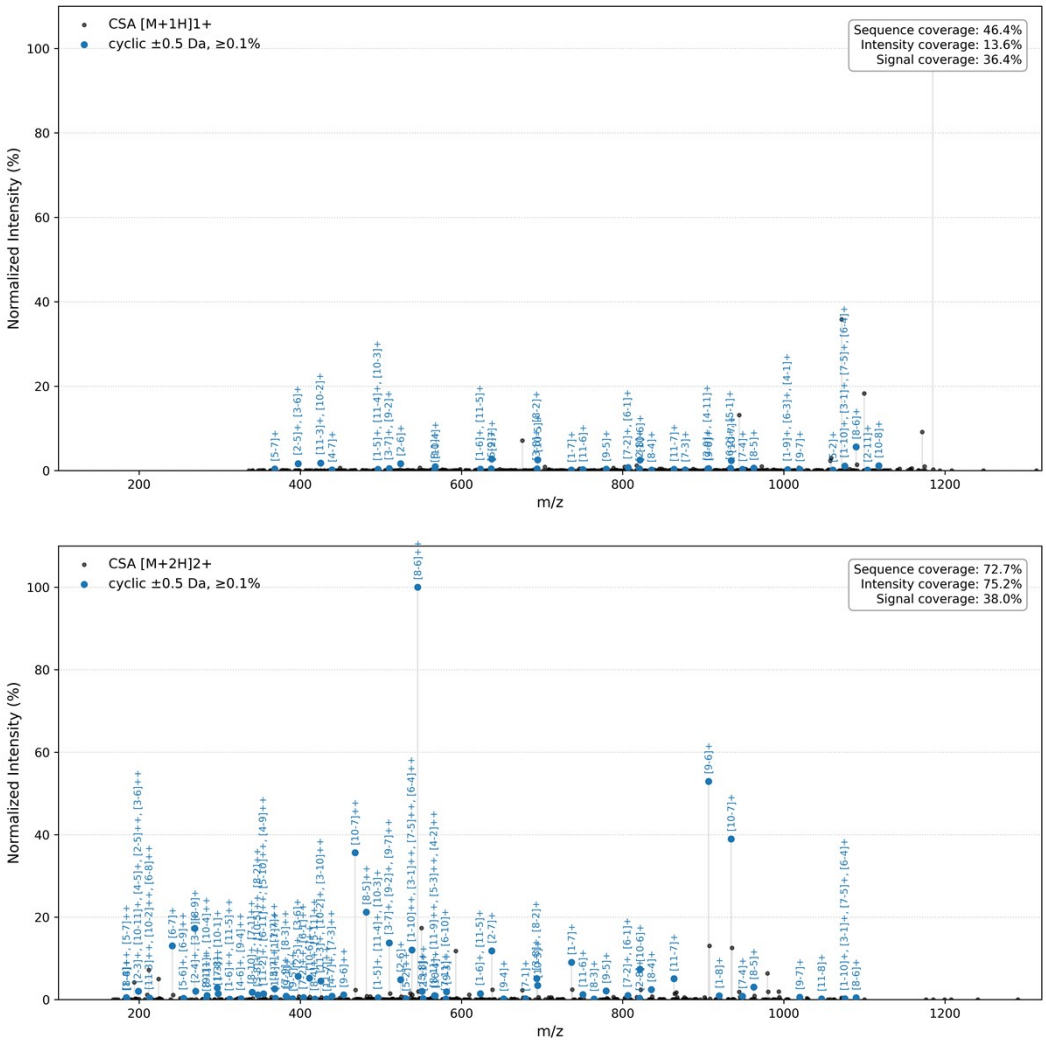


Figure S3-7: MS/MS data assigned CSA

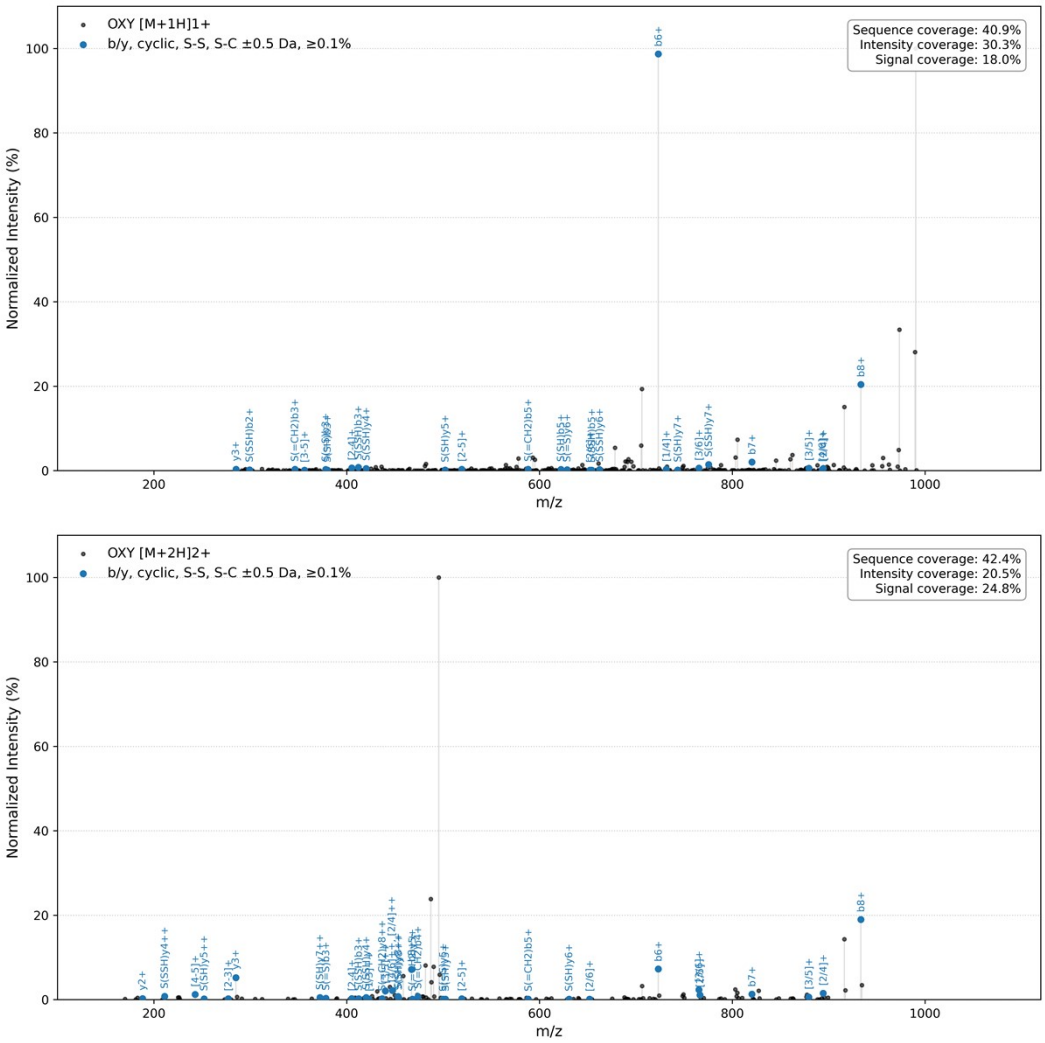


Figure S3-8: MS/MS data assigned OXY

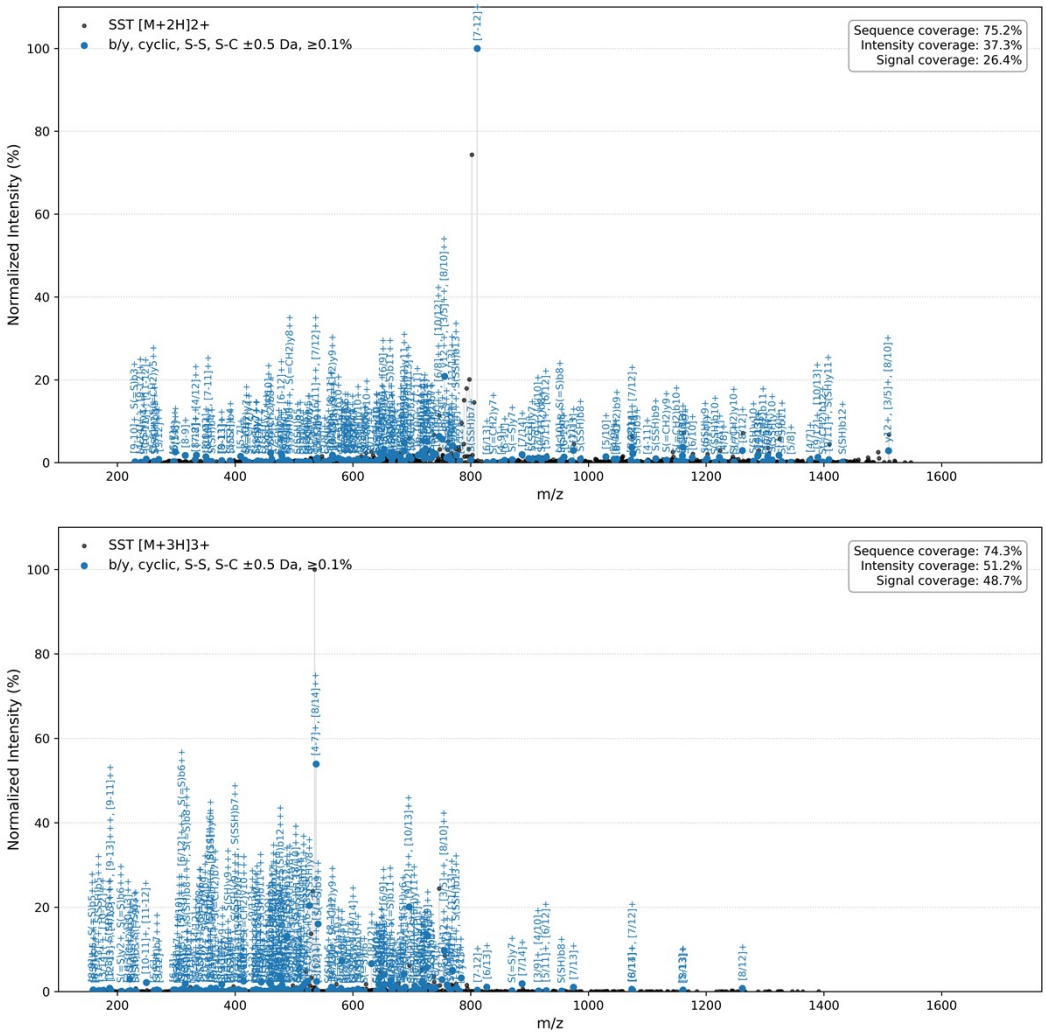


Figure S3-9: MS/MS data assigned SMT

S4 Additional figures

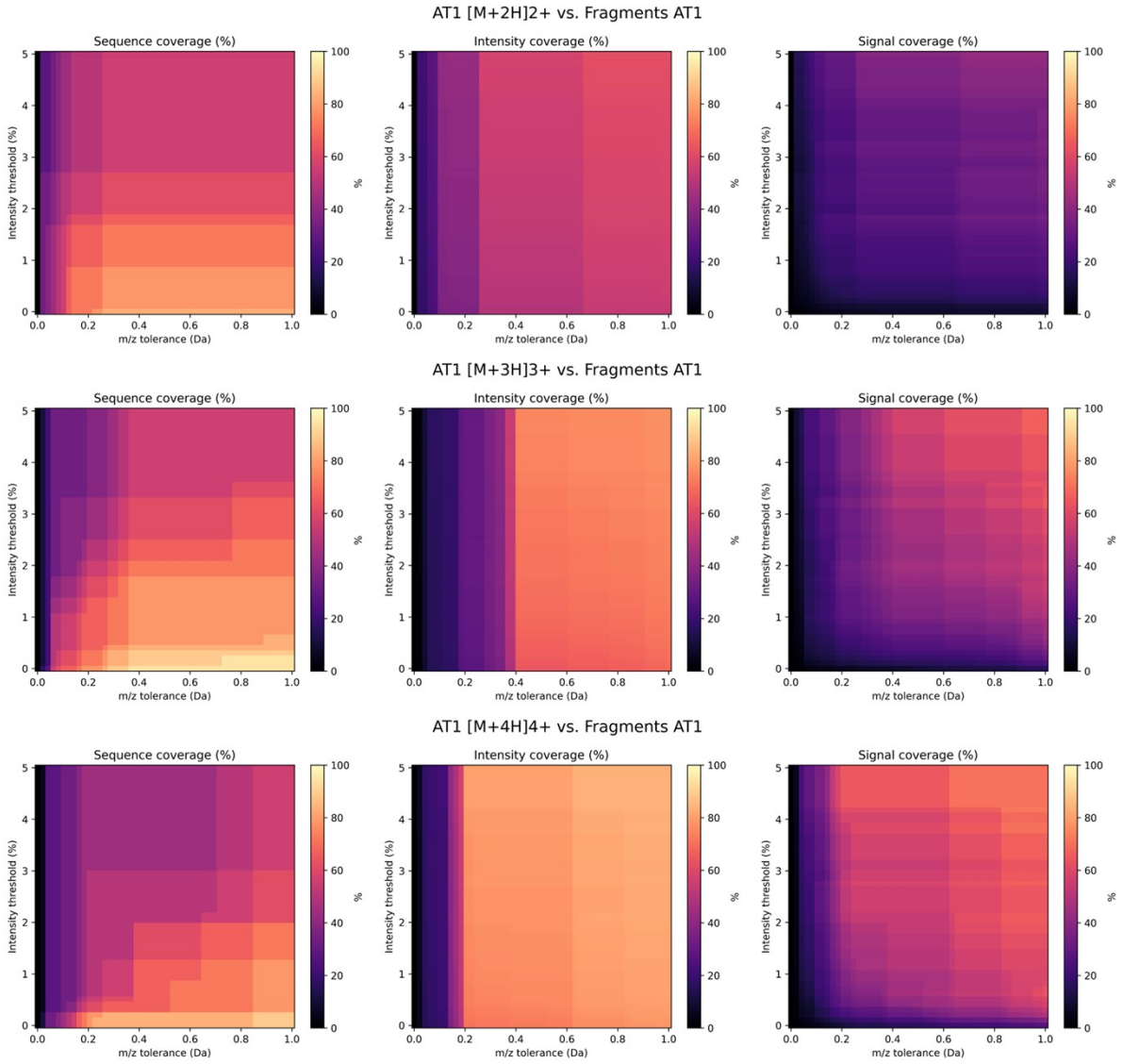


Figure S4-1: Heatmaps showing the sequence coverage, intesity coverage and signal coverage in dependence of the intensity threshold and the m/z threshold defined for all charged states of the AT1 MSMS spectra assigmend with AT1 fragments.

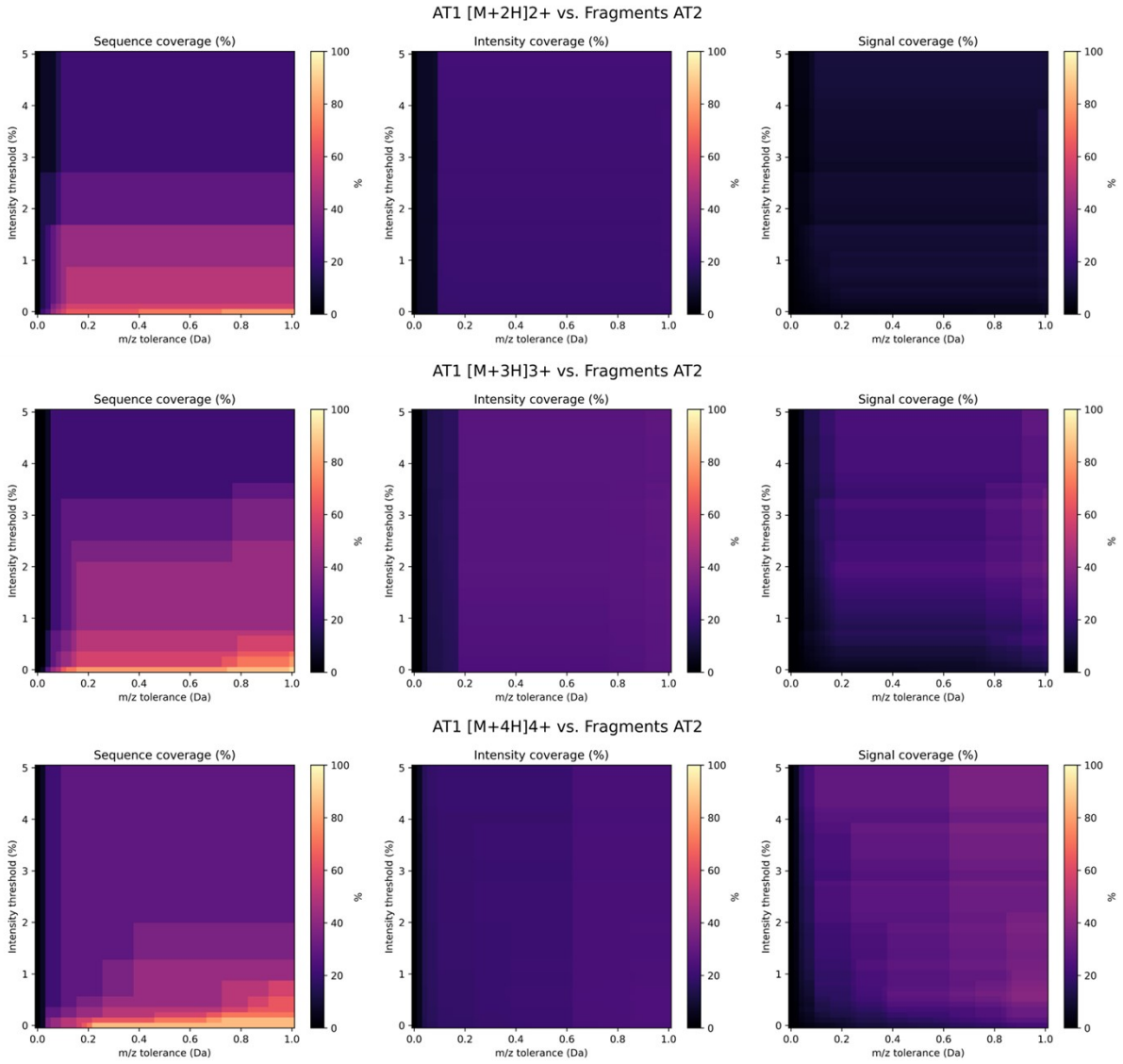


Figure S4-2: Heatmaps showing the sequence coverage, intesity coverage and signal coverage in dependence of the intensity threshold and the m/z threshold defined for all charged states of the AT1 MSMS spectra assigmend with AT2 fragments (related peptide).

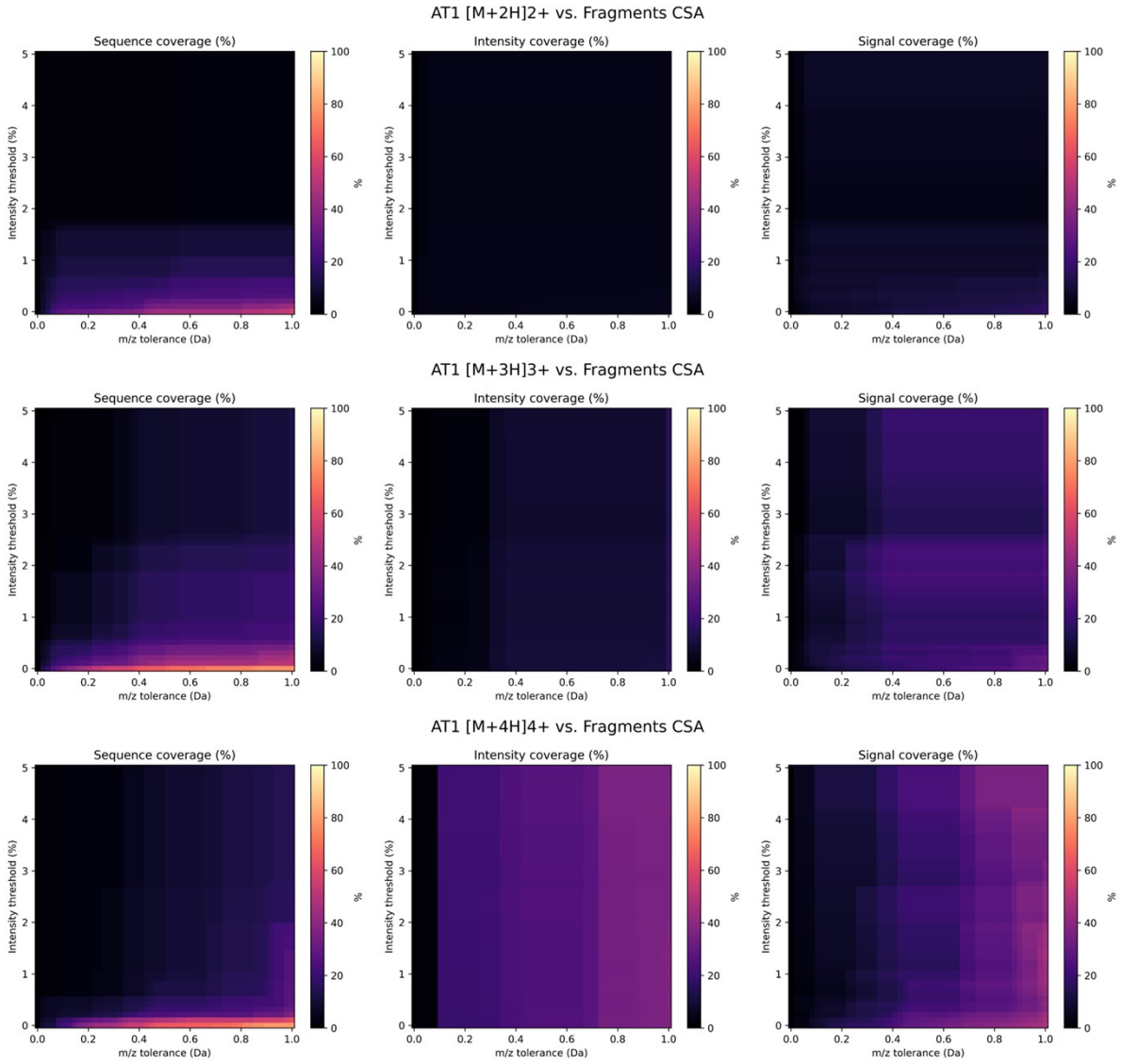


Figure S4-3: Heatmaps showing the sequence coverage, intesity coverage and signal coverage in dependence of the intensity threshold and the m/z threshold defined for all charged states of the AT1 MSMS spectra assigmend with CSA fragments (unrelated peptide in similar mass range with high number of possible fragments).

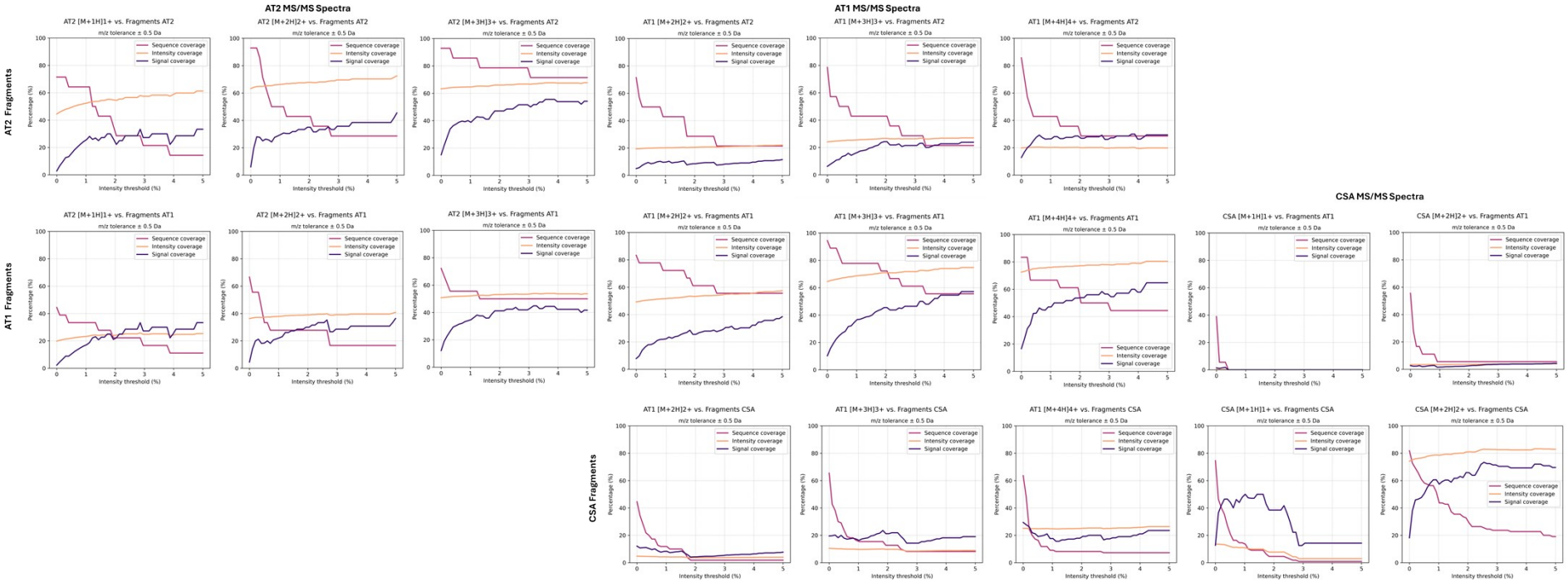


Figure S4-5: Comparison of correct and false MS/MS spectra assigment for linear peptides AT1 and AT2 and head-to-tail cyclized peptide CSA.

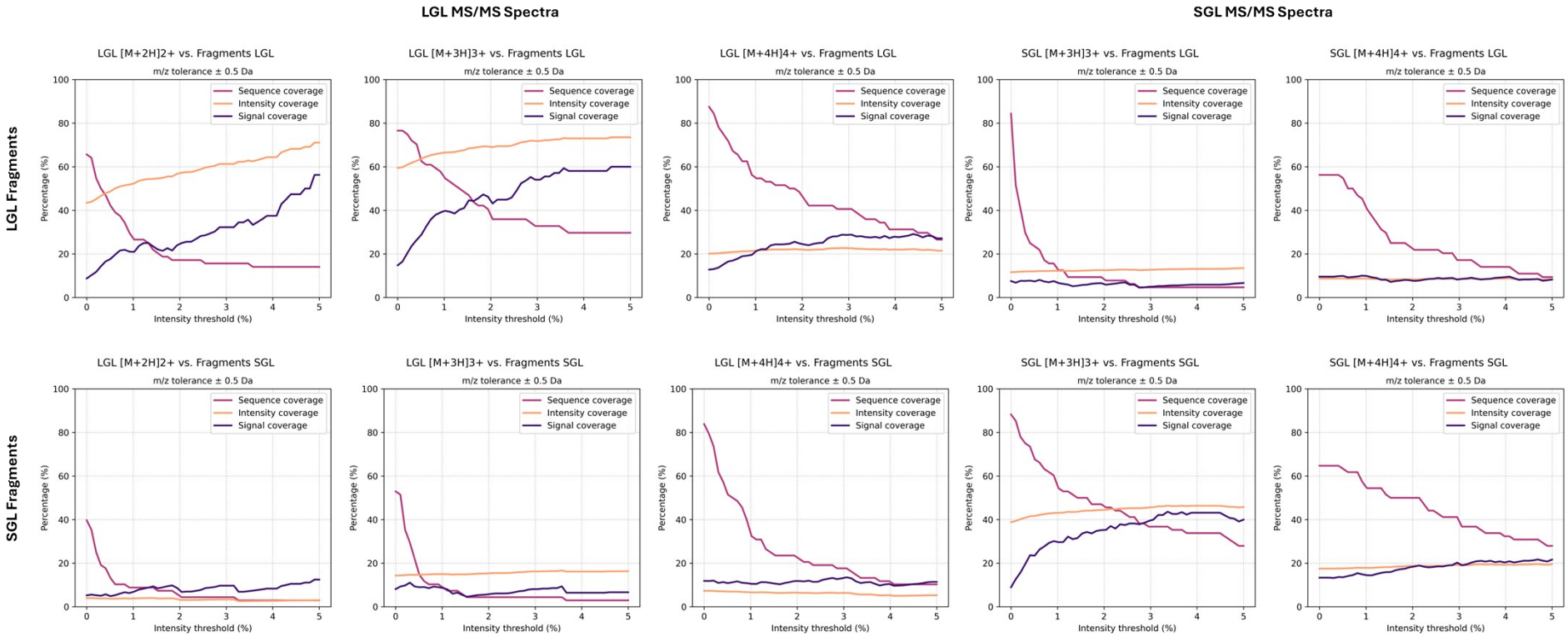


Figure S4-6: Comparison of correct and false MS/MS spectra assigment for fatty acid modified peptides LGL and SGL.

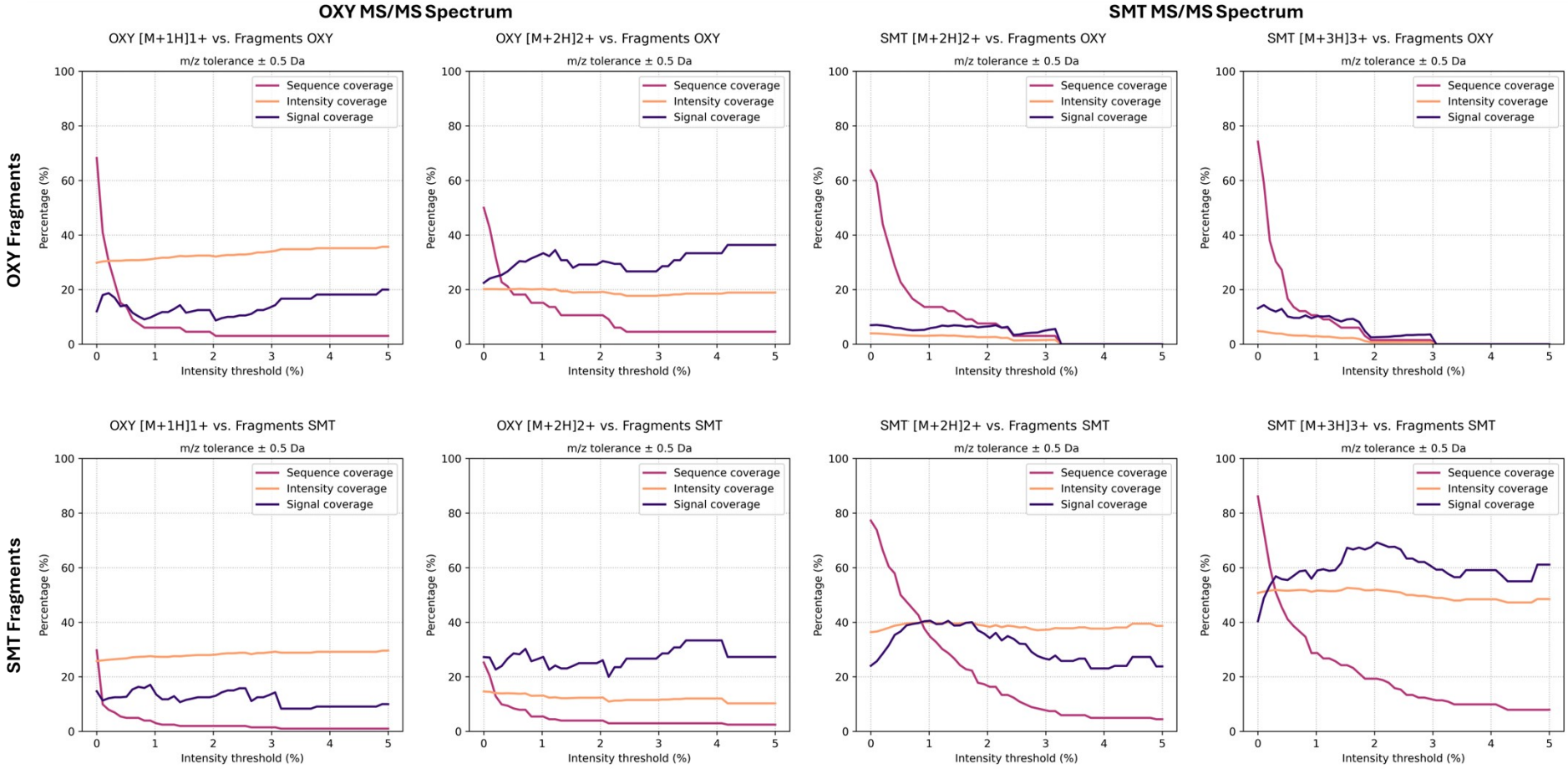


Figure S4-7: Comparison of correct and false MS/MS spectra assigment for peptides OXY and SMT with a disulfide bridge.

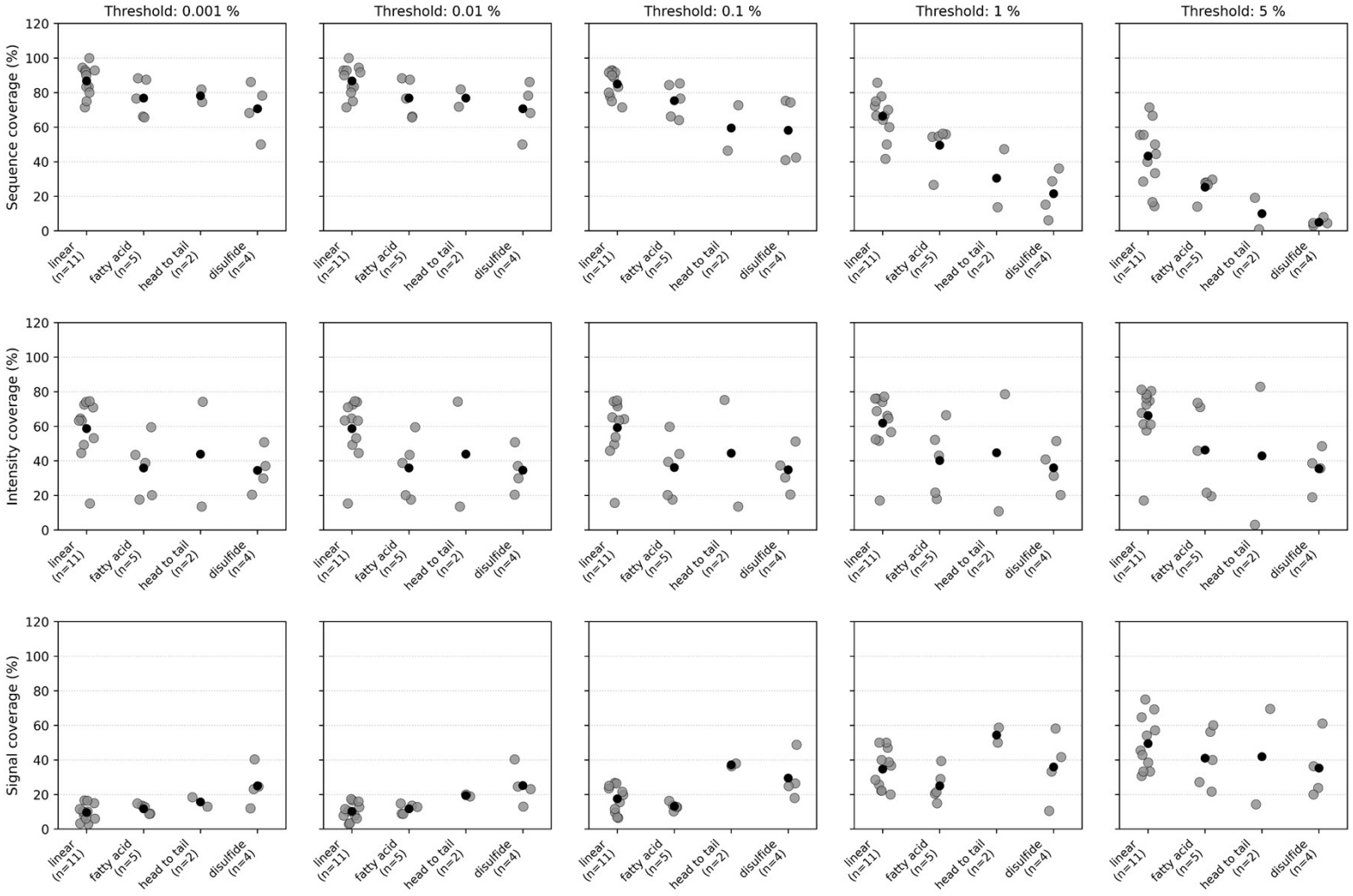


Figure S4-8: Average sequence coverage, intensity coverage and signal coverage of different categories of peptides (linear, fatty acid modified, head-to-tail cyclized and with disulfide bridge) for different intensity thresholds and a m/z tolerance ±0.5 Da defined for assigment. Individual data points are shown in grey and the mean value in black.

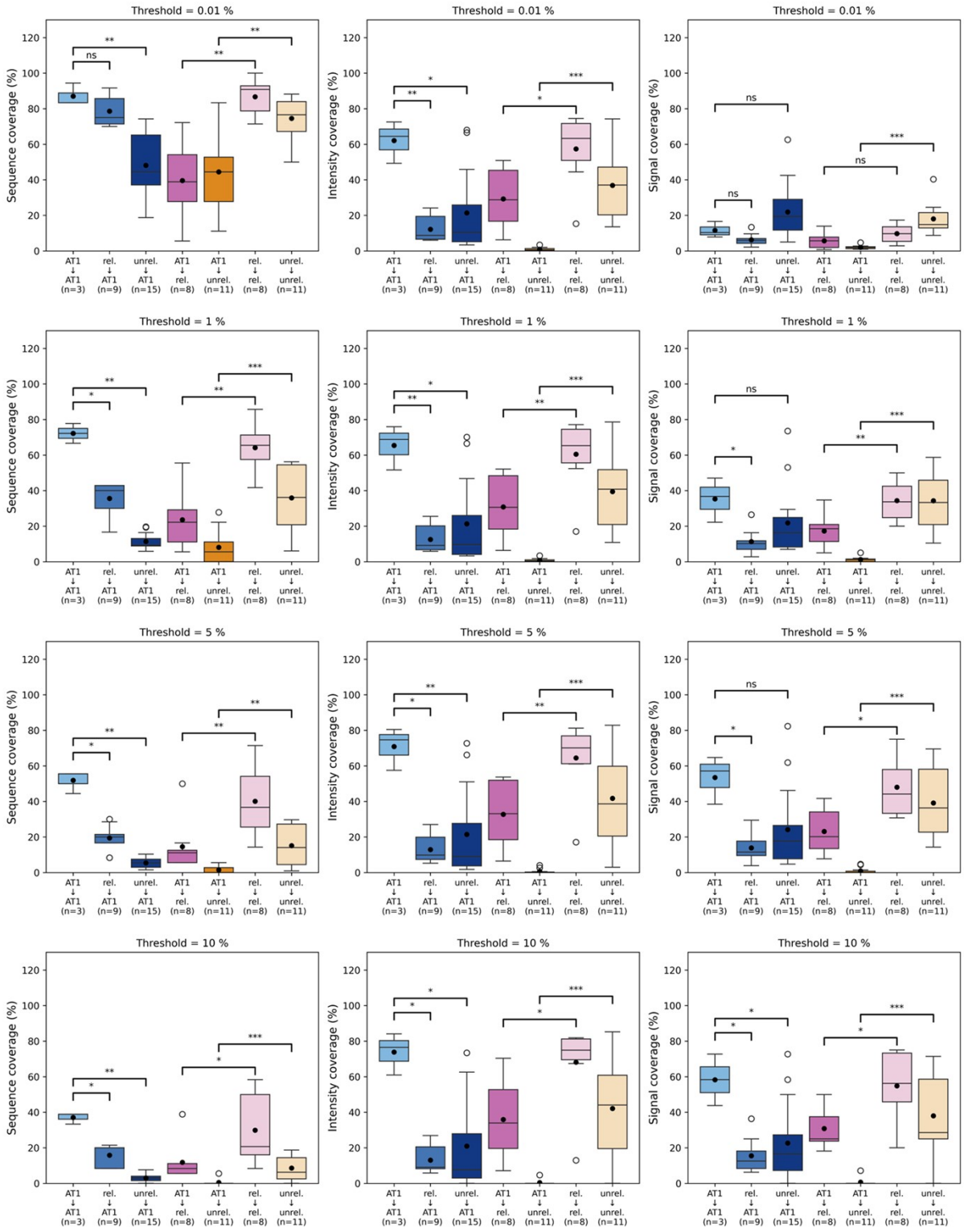


Figure S4-9: Comparison of different assigment metrics to distinguish related and unrelated false assigned spectra for different intensity thresholds (%).

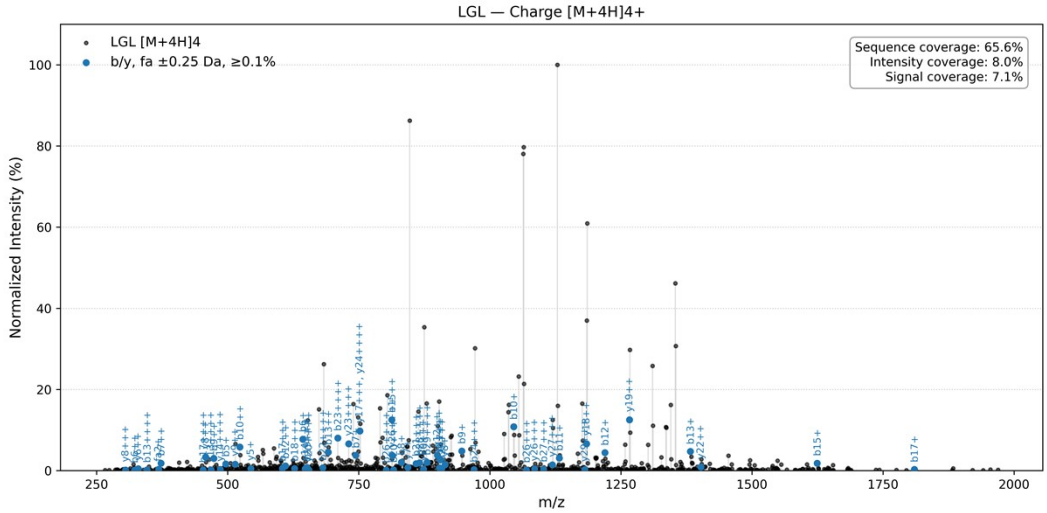


Figure S4-10: LGL $[M+4H]^{4+}$ MS/MS data assigment with a ±0.25 Da tolerance including fa fragments.

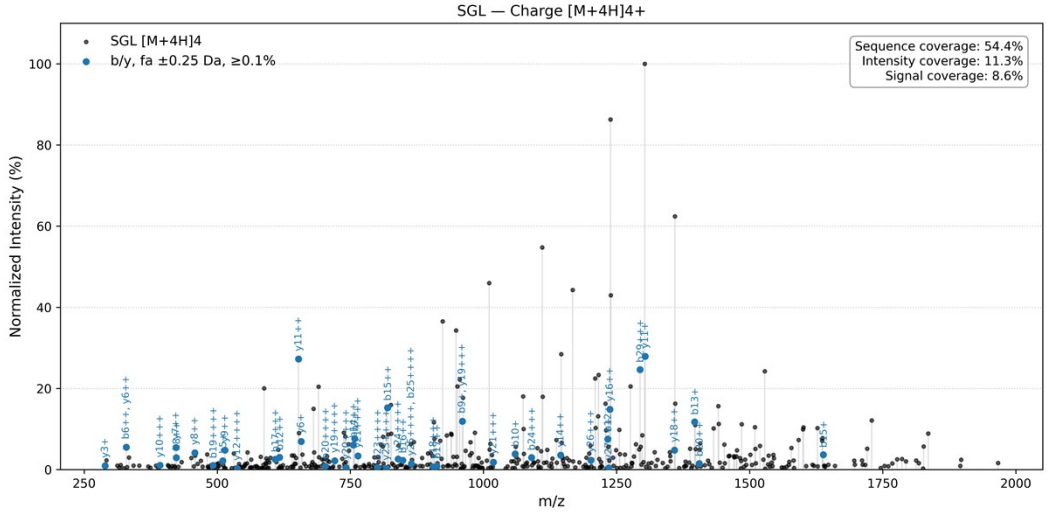


Figure S4-11: SGL $[M+4H]^{4+}$ MS/MS data assigment with a ±0.25 Da tolerance including fa fragments.

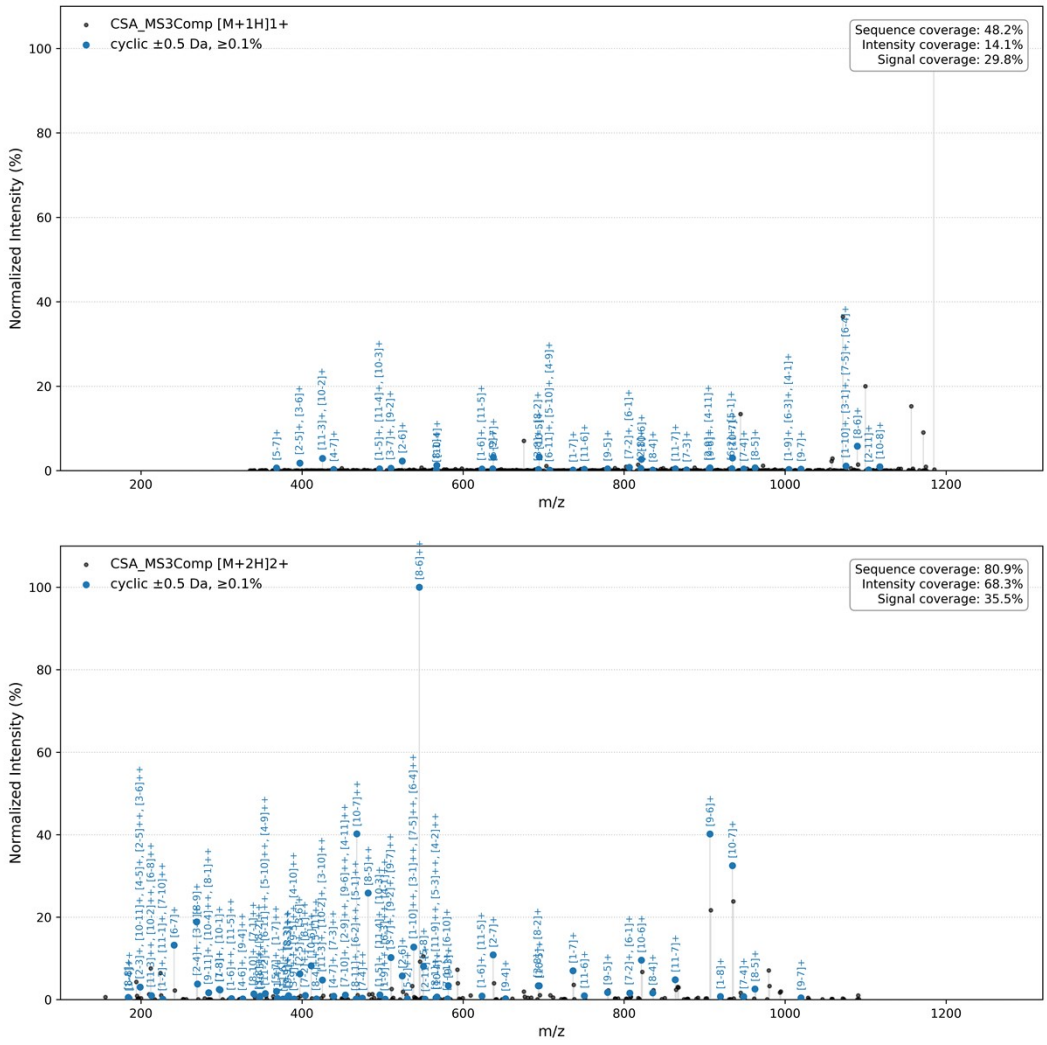


Figure S4-12: MS^3 data assigned CSA

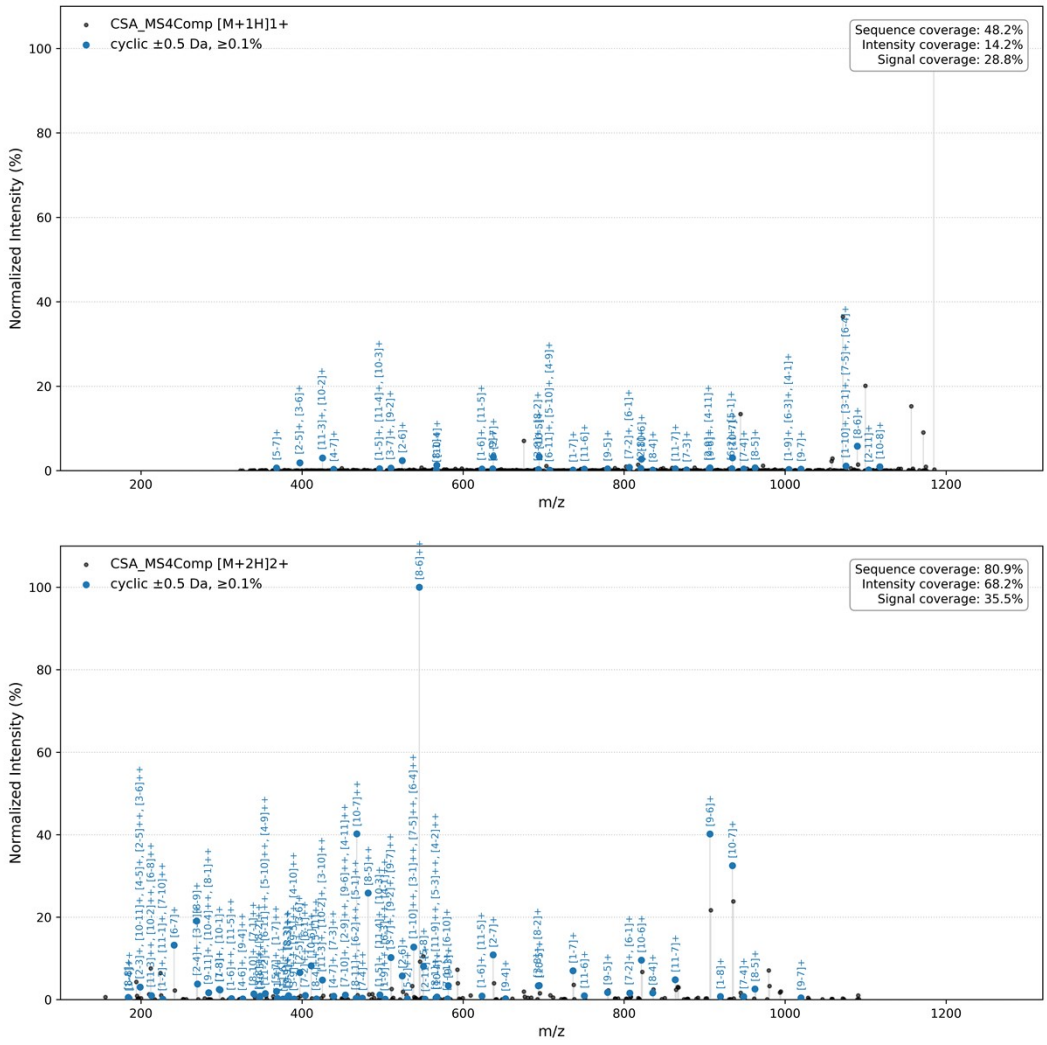


Figure S4-13: MS^4 data assigned CSA

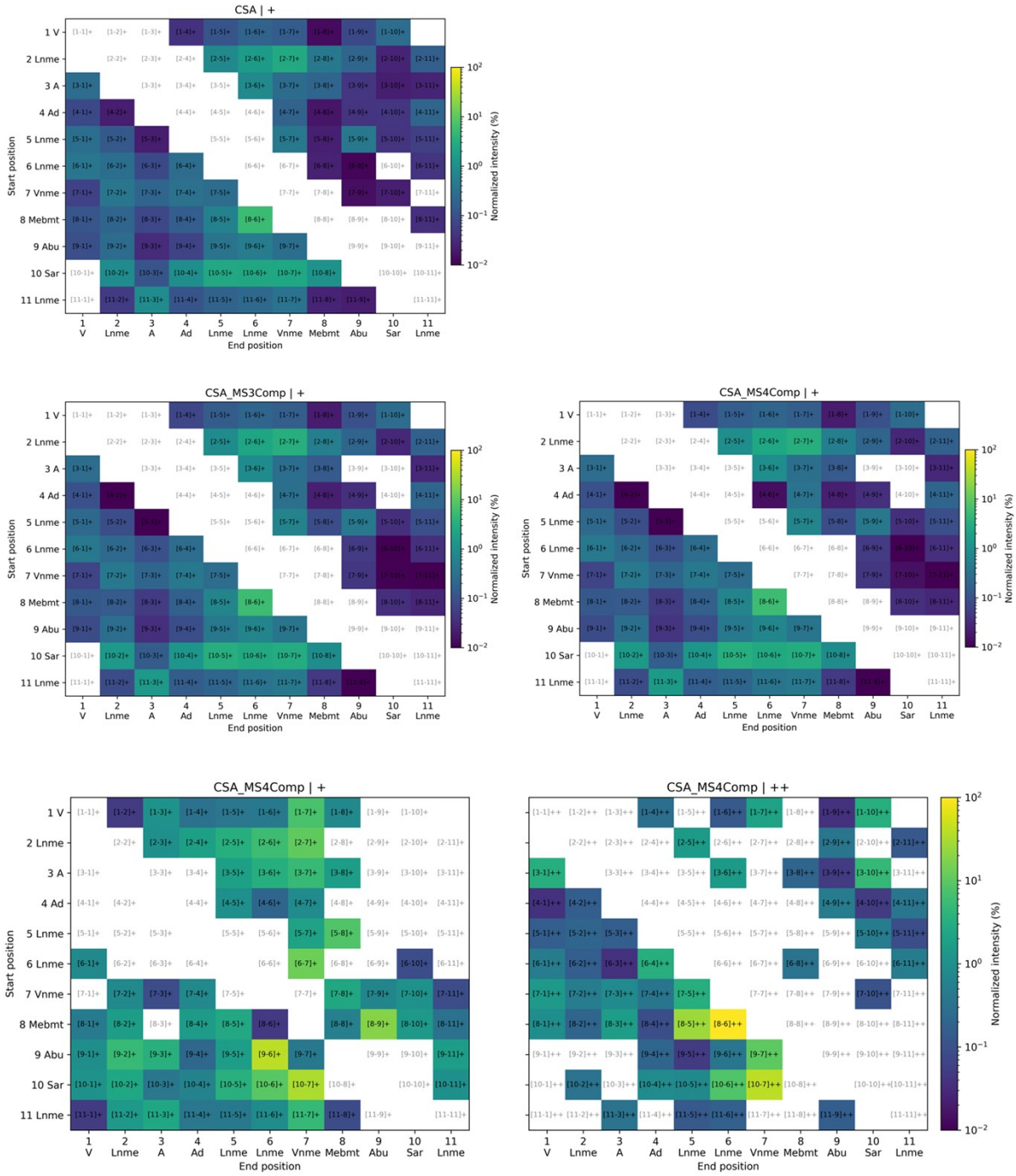


Figure S4-14: Additional data of MS^n analysis of CSA

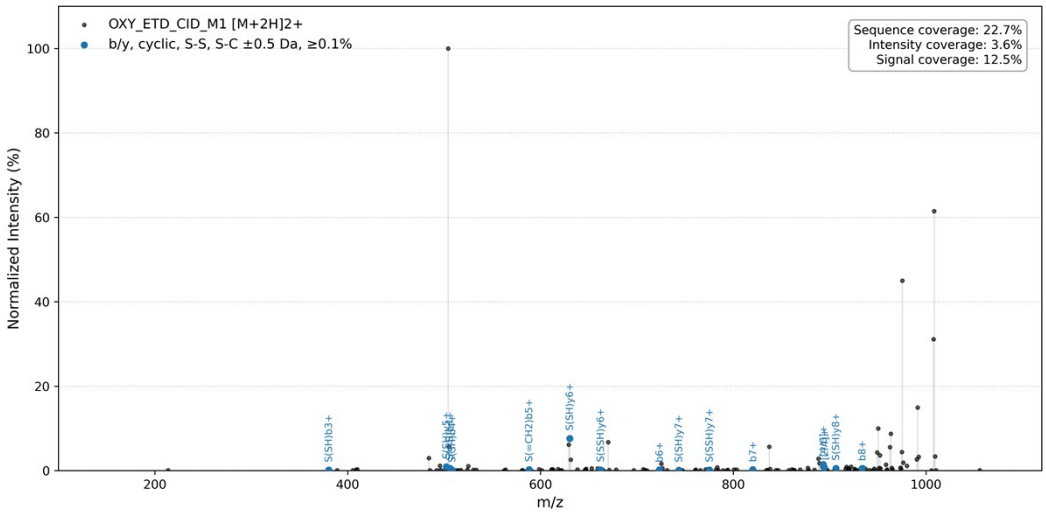


Figure S4-15: ETD-CID data assigned OXY.

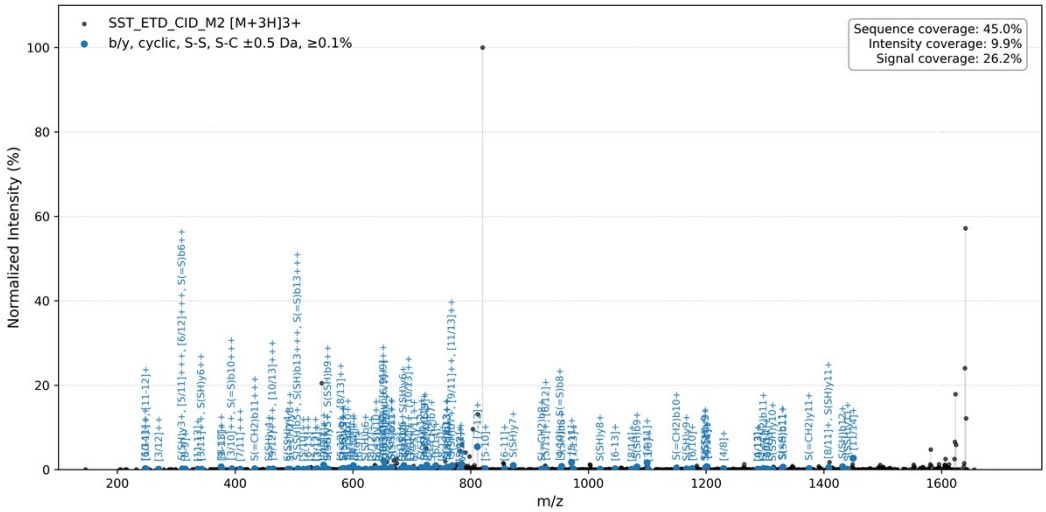


Figure S4-16: ETD-CID data assigned SMT.

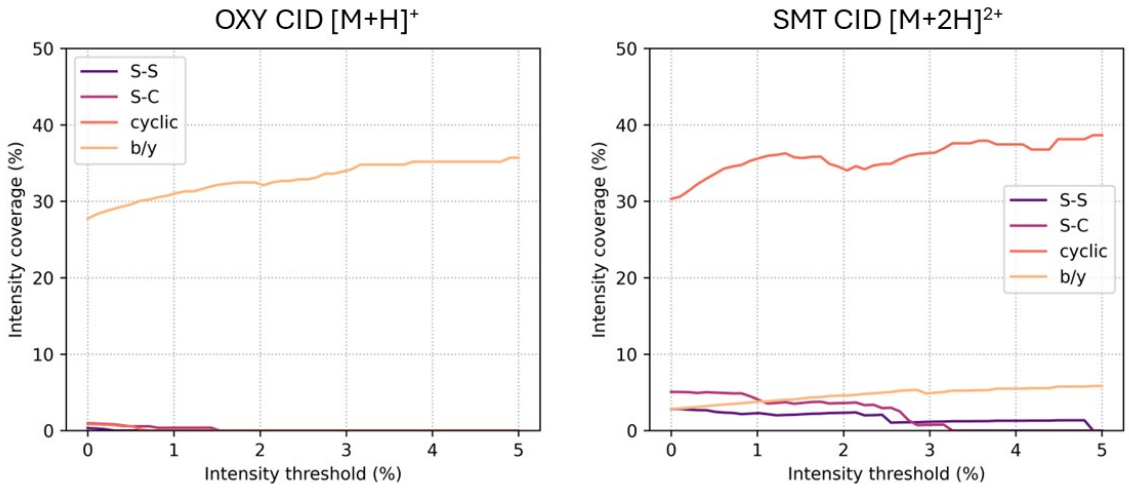


Figure S4-17: Additional charged states for CID fragmentation with intensity coverage by type of fragmentation

S5 References

(1) Egelund, P. H. G.; Jadhav, S.; Martin, V.; Johansson Castro, H.; Richner, F.; Le Quement, S. T.; Dettner, F.; Lechner, C.; Schoenleber, R.; Sejer Pedersen, D. Fmoc-Removal with Pyrrolidine Expands the Available Solvent Space in Green Solid-Phase Peptide Synthesis. *ACS Sustainable Chemistry & Engineering* **2021**, *9* (42), 14202–14215.
(2) Hood, C. A.; Fuentes, G.; Patel, H.; Page, K.; Menakuru, M.; Park, J. H. Fast conventional Fmoc solid-phase peptide synthesis with HCTU. *J Pept Sci* **2008**, *14* (1), 97–101.
(3) Bacsa, B.; Horváti, K.; Bõsze, S.; Andreae, F.; Kappe, C. O. Solid-Phase Synthesis of Difficult Peptide Sequences at Elevated Temperatures: A Critical Comparison of Microwave and Conventional Heating Technologies. *The Journal of Organic Chemistry* **2008**, *73* (19), 7532–7542.